PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, 15/74, 5/10, C07K 14/47, 16/18, A01K 67/027, C12Q 1/68, G01N 33/577, A61K 31/70, 38/17, 39/395, 48/00

A1

(11) International Publication Number: WO 97/27296

(43) International Publication Date: 31 July 1997 (31.07.97)

(21) International Application Number: PCT/CA97/00051

(22) International Filing Date: 27 January 1997 (27.01.97)

(30) Priority Data:

08/592,541	26 January 1996 (26.01.96)	US
60/021,673	5 July 1996 (05.07.96)	US
60/021,700	12 July 1996 (12.07.96)	US
60/029,895	8 November 1996 (08.11.96)	US
60/034,590	2 January 1997 (02.01.97)	US

(71) Applicants *(for all designated States except US)*: HSC RESEARCH AND DEVELOPMENT LIMITED PARTNERSHIP [CA/CA]; Suite 5270, 555 University Avenue, Toronto, Ontario M5G 1X8 (CA). THE GOVERNING COUNCIL OF THE UNIVERSITY OF TORONTO [CA/CA]; 106 Simcoe Hall, 27 King's College Circle, Toronto, Ontario M5S 1A1 (CA).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: ST. GEORGE-HYSLOP, Peter, H. [CA/CA]; 210 Richview Avenue, Toronto, Ontario M5P 3G3 (CA). FRASER, Paul, E. [CA/CA]; 611 Windermere Avenue, Toronto, Ontario M6S 3L9 (CA). ROMMENS, Johanna, M. [CA/CA]; 105 McCaul Street, Toronto, Ontario M5T 2X4 (CA).

(74) Agent: RAE, Patricia, A.; Sim & McBurney, 6th floor, 330 University Avenue, Toronto, Ontario M5G 1R7 (CA).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: NUCLEIC ACIDS AND PROTEINS RELATED TO ALZHEIMER'S DISEASE, AND USES THEREFOR

(57) Abstract

The present invention describes the identification, isolation, sequencing and characterization of several human genes which interact with the presenilins, mutations in which may lead to Familial Alzheimer's Disease. These presenilin-interacting protein genes may be involved in the pathways which, when affected by mutant presenilins, lead to the development of Alzheimer's Disease. In addition, mutations in the presenilin-interacting protein genes, even in the absence of defects in the presenilins, may be causative of Alzheimer's Disease. Nucleic acids and proteins comprising or derived from the presenilin-interacting proteins are useful in screening and diagnosing Alzheimer's Disease, in identifying and developing therapeutics for treatment of Alzheimer's Disease, and in producing cell lines and transgenic animals useful as models of Alzheimer's Disease.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia
AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

NUCLEIC ACIDS AND PROTEINS RELATED TO ALZHEIMER'S DISEASE, AND USES THEREFOR

Field of the Invention

The present invention relates generally to the field of neurological and physiological dysfunctions associated with Alzheimer's Disease. More particularly, the invention is concerned with the identification, isolation and cloning of genes which are associated with Alzheimer's Disease, as well as their corresponding transcripts and protein products. The present invention also relates to methods for detecting and diagnosing carriers of normal and mutant alleles of these genes, to methods for detecting and diagnosing Alzheimer's Disease, to methods of identifying other genes and proteins related to, or interacting with, the genes and proteins of the invention, to methods of screening for potential therapeutics for Alzheimer's Disease, to methods of treatment for Alzheimer's Disease, and to cell lines and animal models useful in screening for and evaluating potentially useful therapies for Alzheimer's Disease.

Background of the Invention

Alzheimer's Disease (AD) is a degenerative disorder of the human central nervous system characterized by progressive memory impairment and cognitive and intellectual decline during mid to late adult life (Katzman, 1986). The disease is accompanied by a constellation of neuro-pathologic features principal amongst which are the presence of extracellular amyloid or senile plaques, and neurofibrillary tangles in neurons. The etiology of this disease is complex, although in some families it appears to be inherited as an autosomal dominant trait. Linkage studies have identified three genes associated with the development of AD: β-amyloid precursor protein (APP) (Chartier-Harlin et al., 1991; Goate et al., 1991; Murrell et al., 1991; Karlinsky et al., 1992; Mullan et al., 1992), presenilin-1 (PS-1) (Sherrington, 1995), and presenilin-2 (PS-2) (Rogaev, 1995, and Levy-Lahad, 1995).

The presenilins are multi-spanning membrane proteins which were described in substantial detail in PCT Publication WO96/34099, the entire disclosure of which is incorporated herein by reference. Although the functions of the presenilins are unknown, a number of autosomal dominant presenilin mutations have been identified

which are strongly associated with the development of early-onset, aggressive, Familial Alzheimer's Disease (FAD).

The present disclosure describes the identification, isolation, sequencing and characterization of several human genes which interact with the presenilins, mutations in which may lead to FAD. These presenilin-interacting protein genes may be involved in the pathways which, when affected by mutant presenilins, lead to the development of Alzheimer's Disease. In addition, mutations in the presenilin-interacting protein genes, even in the absence of defects in the presenilins, may be causative of Alzheimer's Disease.

Summary of the Invention

The present invention is based, in part, upon the identification, isolation, sequencing and characterization of several human genes, referred to herein as "presenilin-interacting protein genes" or "PS-interacting protein genes." The products of these genes are believed to interact in vivo with the human presenilin-1 proteins and, therefore, are implicated in the biochemical pathways which are affected in Alzheimer's Disease. Each of these genes, therefore, presents a new therapeutic target for the treatment of Alzheimer's Disease. In addition, PS-interacting protein nucleic acids, PS-interacting proteins and peptides, antibodies to the PS-interacting proteins, cells transformed with PS-interacting protein nucleic acids, and transgenic animals altered with PS-interacting protein nucleic acids, all possess various utilities, as described herein, for the diagnosis, therapy and continued investigation of Alzheimer's Disease and related disorders.

Thus, it is one object of the invention to provide isolated nucleic acids encoding at least a PS-interacting domain of a PS-interacting protein. These PS-interacting proteins include mammalian S5a subunits of the 26S proteasome, the GT24 protein, the p0071 protein, the Rab11 protein, the retinoid X receptor-β, the cytoplasmic chaperonin, and several sequences identified herein as clones Y2H35, Y2H171, and Y2H41. Preferred nucleotide and amino acid sequences are provided herein. It is another object of the invention to provide probes and primers for these

PS-interacting protein genes, and to provide nucleic acids which encode small antigenic determinants of these genes. Therefore, preferred embodiments include sequences of at least 10, 15 or 20 consecutive nucleotides selected from the disclosed sequences.

Using the nucleic acid sequences and antibodies disclosed and enabled herein, methods for identifying allelic variants or heterospecific homologues of a human PS-interacting protein and gene are provided. The methods may be practiced using nucleic acid hybridization or amplification techniques, immunochemical techniques, or any other technique known in the art. The allelic variants may include other normal human alleles as well as mutant alleles of the PS-interacting protein genes which may be causative of Alzheimer's Disease. The heterospecific homologues may be from other mammalian species, such as mice, rats, dogs, cats or non-human primates, or may be from invertebrate species, such as Drosophila or C. elegans. Thus, it is another object of the invention to provide nucleic acids which encode allelic or heterospecific variants of the disclosed sequences, as well as the allelic or heterospecific proteins encoded by them.

The it another object of the invention to provide vectors, and particularly expression vectors, which include any of the above-described nucleic acids. It is a further object of the invention to provide vectors in which PS-interacting protein nucleic acid sequences are operably joined to exogenous regulatory regions to produce altered patterns of expression, or to exogenous coding regions to produce fusion proteins. Conversely, it is another object to provide nucleic acids in which PS-interacting protein regulatory regions are operably joined to exogenous coding regions, including standard marker genes, to produce constructs in which the regulation of PS-interacting protein genes may be studied and used in assays for therapeutics.

It is another object of the invention to provide host cells and transgenic animals which have been transformed with any of the above-described nucleic acids

of the invention. The host cells may be prokaryotic or eukaryotic cells and, in particular, may be gametes, zygotes, fetal cells, or stem cells useful in producing transgenic animal models.

In particularly preferred embodiments, the present invention provides a non-human animal model for Alzheimer's Disease, in which the genome of the animal, or an ancestor thereof, has been modified by at least one recombinant construct which has introduced one of the following modifications: (1) insertion of nucleotide sequences encoding at least a functional domain of a heterospecific normal PS-interacting protein, (2) insertion of nucleotide sequences encoding at least a functional domain of a heterospecific mutant PS-interacting protein, (3) insertion of nucleotide sequences encoding at least a functional domain of a conspecific homologue of a heterospecific mutant PS-interacting protein, and (4) inactivation of an endogenous PS-interacting protein gene. Preferred transgenic animal models are rats, mice, hamsters, guinea pigs, rabbits, dogs, cats, goats, sheep, pigs, and non-human primates, but invertebrates are also contemplated for certain utilities.

It is another object of the invention to provide methods for producing at least a functional domain of a PS-interacting protein using the nucleic acids of the invention. In addition, the present invention also provides substantially pure preparations of such proteins, including short peptide sequences for used as immunogens. Thus, the invention provides peptides comprising at least 10 or 15 consecutive amino acid residues from the disclosed and otherwise enabled sequences. The invention further provides substantially pure preparations of peptides which comprise at least a PS-interacting domain of a PS-interacting protein, as well as substantially pure preparations of the entire proteins.

Using the substantially pure peptides and proteins enabled herein, the invention also provides methods for producing antibodies which selectively bind to a PS-interacting protein, as well as cell lines which produce these antibodies.

Another object of the present invention is to provide methods of identifying compounds which may have utility in the treatment of Alzheimer's Disease and related disorders. These methods include methods for identifying compounds which can modulate the expression of a PS-interacting protein gene, methods for identifying compounds which can selectively bind to a PS-interacting protein, and methods of identifying compounds which can modulate activity of a PS-interacting protein. These methods may be conducted in vitro or in vivo, and may employ the transformed cell lines and transgenic animal models of the invention. The methods also may be part of a clinical trial in which compounds identified by the methods of the invention are further tested in human subjects.

It is another object of the invention to provide methods of diagnosing or screening for inherited forms of Alzheimer's Disease by determining if a subject bears a mutant PS-interacting protein gene. Mutant PS-interacting genes may be detected by assays including direct nucleotide sequencing, probe specific hybridization, restriction enzyme digest and mapping, PCR mapping, ligase-mediated PCR detection, RNase protection, electrophoretic mobility shift detection, or chemical mismatch cleavage. Alternatively, mutant forms of a PS-interacting protein may be detected by assays including immunoassays, protease assays, or electrophoretic mobility assays.

It is also an object of the invention to provide pharmaceutical preparations which may be used in the treatment of Alzheimer's Disease and related disorders which result from aberration in biochemical pathways involving the PS-interacting proteins disclosed and enabled herein. Thus, the present invention also provides pharmaceutical preparations comprising a substantially pure PS-interacting protein, an expression vector operably encoding a PS -interacting protein, an expression vector operably encoding a PS-interacting protein antisense sequence, an antibody which selectively binds to a mutant PS-interacting protein, or an antigenic determinant of a mutant PS-interacting protein. These pharmaceutical preparations may be used to

treat a patient bearing a mutant PS-interacting protein gene which is causative of Alzheimer's Disease or related disorders.

These an other objects of the present invention are described more fully in the following specification and appended claims.

Detailed Description of the Invention

I. Definitions

In order to facilitate review of the various embodiments of the invention, and an understanding of the various elements and constituents used in making and using the invention, the following definitions are provided for particular terms used in the description and appended claims:

Presenilin. As used without further modification herein, the terms "presenilin" or "presenilins" mean the presenilin-1 (PS1) and/or the presenilin-2 (PS2) genes/proteins. In particular, the unmodified terms "presenilin" or "presenilins" refer to the mammalian PS1 and/or PS2 genes/proteins and, preferably, the human PS1 and/or PS2 genes/proteins as described and disclosed in PCT Publication WO96/34099.

Normal. As used herein with respect to genes, the term "normal" refers to a gene which encodes a normal protein. As used herein with respect to proteins, the term "normal" means a protein which performs its usual or normal physiological role and which is not associated with, or causative of, a pathogenic condition or state. Therefore, as used herein, the term "normal" is essentially synonymous with the usual meaning of the phrase "wild type." For any given gene, or corresponding protein, a multiplicity of normal allelic variants may exist, none of which is associated with the development of a pathogenic condition or state. Such normal allelic variants include, but are not limited to, variants in which one or more nucleotide substitutions do not result in a change in the encoded amino acid sequence.

Mutant. As used herein with respect to genes, the term "mutant" refers to a gene which encodes a mutant protein. As used herein with respect to proteins, the term "mutant" means a protein which does not perform its usual or normal

physiological role and which is associated with, or causative of, a pathogenic condition or state. Therefore, as used herein, the term "mutant" is essentially synonymous with the terms "dysfunctional," "pathogenic," "disease-causing," and "deleterious." With respect to the presenilin and presenilin-interacting protein genes and proteins of the present invention, the term "mutant" refers to genes/proteins bearing one or more nucleotide/amino acid substitutions, insertions and/or deletions which typically lead to the development of the symptoms of Alzheimer's Disease and/or other relevant inheritable phenotypes (e.g. cerebral hemorrhage, mental retardation, schizophrenia, psychosis, and depression) when expressed in humans. This definition is understood to include the various mutations that naturally exist, including but not limited to those disclosed herein, as well as synthetic or recombinant mutations produced by human intervention. The term "mutant," as applied to these genes, is not intended to embrace sequence variants which, due to the degeneracy of the genetic code, encode proteins identical to the normal sequences disclosed or otherwise enabled herein; nor is it intended to embrace sequence variants which, although they encode different proteins, encode proteins which are functionally equivalent to normal proteins.

Substantially pure. As used herein with respect to proteins (including antibodies) or other preparations, the term "substantially pure" means that the preparation is essentially free of other substances to an extent practical and appropriate for its intended use. In particular, a protein preparation is substantially pure if it is sufficiently free from other biological constituents so as to be useful in, for example, generating antibodies, sequencing, or producing pharmaceutical preparations. By techniques well known in the art, substantially pure proteins or peptides may be produced in light of the nucleic acid and amino acid sequences disclosed herein. In particular, in light of the nucleic acid and amino acid sequences disclosed herein, one of ordinary skill in the art may, by application or serial application of well-known methods including HPLC or immuno-affinity chromatography or electrophoretic separation, obtain proteins or peptides of any generally feasible purity. Preferably, but not necessarily, "substantially pure"

preparations include at least 60% by weight (dry weight) the compound of interest. More preferably the preparation is at least 75% or 90%, and most preferably at least 99%, by weight the compound of interest. Purity can be measured by any appropriate method, e.g., column chromatography, gel electrophoresis, or HPLC analysis. With respect to proteins, including antibodies, if a preparation includes two or more different compounds of interest (e.g., two or more different antibodies, immunogens, functional domains, or other polypeptides of the invention), a "substantially pure" preparation is preferably one in which the total weight (dry weight) of all the compounds of interest is at least 60% of the total dry weight. Similarly, for such preparations containing two or more compounds of interest, it is preferred that the total weight of the compounds of interest be at least 75%, more preferably at least 90%, and most preferably at least 99%, of the total dry weight of the preparation. Finally, in the event that the protein of interest is mixed with one or more other proteins (e.g., serum albumin) or compounds (e.g., diluents, excipients, salts, polysaccharides, sugars, lipids) for purposes of administration, stability, storage, and the like, such other proteins or compounds may be ignored in calculation of the purity of the preparation.

Isolated nucleic acid. As used herein, an "isolated nucleic acid" is a ribonucleic acid, deoxyribonucleic acid, or nucleic acid analog comprising a polynucleotide sequence that is isolated or separate from sequences that are immediately contiguous (one on the 5' end and one on the 3' end) in the naturally occurring genome of the organism from which it is derived. The term therefore includes, for example, a recombinant nucleic acid which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (e.g., a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequences and/or including exogenous regulatory elements.

Substantially identical sequence. As used herein, a "substantially identical" amino acid sequence is an amino acid sequence which differs only by conservative amino acid substitutions, for example, substitution of one amino acid for another of the same class (e.g., valine for glycine, arginine for lysine, etc.) or by one or more non-conservative substitutions, deletions, or insertions located at positions of the amino acid sequence which do not destroy the function of the protein (assayed, e.g., as described herein). Preferably, such a sequence is at least 85%, more preferably 90%, and most preferably 95% identical at the amino acid level to the sequence of the protein or peptide to which it is being compared. For nucleic acids, the length of comparison sequences will generally be at least 50 nucleotides, preferably at least 60 nucleotides, more preferably at least 75 nucleotides, and most preferably 110 nucleotides. A "substantially identical" nucleic acid sequence codes for a substantially identical amino acid sequence as defined above.

Transformed cell. As used herein, a "transformed cell" is a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid molecule of interest. The nucleic acid of interest will typically encode a peptide or protein. The transformed cell may express the sequence of interest or may be used only to propagate the sequence. The term "transformed" may be used herein to embrace any method of introducing exogenous nucleic acids including, but not limited to, transformation, transfection, electroporation, microinjection, viral-mediated transfection, and the like.

Operably joined. As used herein, a coding sequence and a regulatory region are said to be "operably joined" when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory region. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of promoter function results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the regulatory region to direct the transcription of the coding sequences, or (3) interfere with the ability of the

corresponding RNA transcript to be translated into a protein. Thus, a regulatory region would be operably joined to a coding sequence if the regulatory region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

Stringent hybridization conditions. Stringent hybridization conditions is a term of art understood by those of ordinary skill in the art. For any given nucleic acid sequence, stringent hybridization conditions are those conditions of temperature, chaotrophic acids, buffer, and ionic strength which will permit hybridization of that nucleic acid sequence to its complementary sequence and not to substantially different sequences. The exact conditions which constitute "stringent" conditions, depend upon the nature of the nucleic acid sequence, the length of the sequence, and the frequency of occurrence of subsets of that sequence within other non-identical sequences. By varying hybridization conditions from a level of stringency at which non-specific hybridization occurs to a level at which only specific hybridization is observed, one of ordinary skill in the art can, without undue experimentation, determine conditions which will allow a given sequence to hybridize only with complementary sequences. Suitable ranges of such stringency conditions are described in Krause and Aaronson (1991). Hybridization conditions, depending upon the length and commonality of a sequence, may include temperatures of 20°C-65°C and ionic strengths from 5x to 0.1x SSC. Highly stringent hybridization conditions may include temperatures as low as 40-42°C (when denaturants such as formamide are included) or up to 60-65°C in ionic strengths as low as 0.1x SSC. These ranges, however, are only illustrative and, depending upon the nature of the target sequence, and possible future technological developments, may be more stringent than necessary. Less than stringent conditions are employed to isolate nucleic acid sequences which are substantially similar, allelic or homologous to any given sequence.

Selectively binds. As used herein with respect to antibodies, an antibody is said to "selectively bind" to a target if the antibody recognizes and binds the target of interest but does not substantially recognize and bind other molecules in a sample, e.g., a biological sample, which includes the target of interest. That is, the antibody

must bind to its target with sufficient specificity so as to distinguish the target from essentially all of molecules which would reasonably be present in a biological sample including the target.

II. The Presenilins and Presenilin-Interacting Proteins

The present invention is based, in part, upon the discovery of a family of mammalian genes which, when mutated, are associated with the development of Alzheimer's Disease. The discovery of these genes, designated presenilin-1 (PS1) and presenilin-2 (PS2), as well as the characterization of these genes, their protein products, mutants, invertebrate homologues, and possible functional roles, are described in PCT Publication WO96/34099. The present invention is further based, in part, upon the discovery of a group of proteins which interact with the presenilins under physiological conditions and which, therefore, are believed to be involved in the biochemical pathways which are altered in Alzheimer's Disease. These proteins are referred to herein as presenilin-interacting (PS-interacting) proteins. Because mutations in the presenilins are known to be causative of Alzheimer's Disease, each of the PS-interacting genes and proteins disclosed and described herein presents a novel target for therapeutic intervention in Alzheimer's Disease. That is, modulation of the interactions of these proteins with the presenilins, or modulation of the interactions of at least the PS-interacting domains of these PS-interacting proteins with at least the interacting domains of the presenilins, provides a means of modulating the activity and/or availability of the presenilins, or of modulating the activity and/or availability of the PS-interacting proteins. Furthermore, as aberrations in the interactions of mutant presenilins with one or more of these PS-interacting proteins is causative of Alzheimer's Disease, mutations in one or more of these PS-interacting proteins are also likely to be causative of Alzheimer's Disease. Therefore, each of the PS-interacting genes and proteins disclosed and described herein presents a novel target for diagnosis of forms of familial and/or sporadic Alzheimer's Disease with an etiology independent of mutations in the presenilins. Finally, as described more fully below, the PS-interacting genes and proteins described and disclosed herein provide for new assays for compounds which affect the interactions of the presenilins and PS-

interacting proteins, assays for other members of the biochemical pathways involved in the etiology of Alzheimer's Disease, and new cell lines and transgenic animal models for use in such assays.

1. Presenilin Processing

Employing the antibodies and protein-binding assays described and/or enabled in PCT Publication WO96/34099, the processing and protein-protein interactions of both normal and mutant presenilins were investigated. It was found that mutations in the presenilins appear to lead to changes in both their intracellular processing (e.g., endoproteolytic cleavage, ubiquitination, and clearance) and their intracellular interactions with other proteins expressed in human brain. As described below, knowledge of presenilin processing and interactions, and particularly changes in mutant presenilin processing and interactions, provides for new diagnostic and therapeutic targets for Alzheimer's Disease and related disorders.

Western blot analysis suggests that the normal presenilins undergo proteolytic cleavage to yield characteristic N- and C-terminal fragments. As noted above, the normal presenilin proteins have an expected molecular mass of 47-51 kDa depending, in part, upon mRNA splice variations, electrophoretic conditions, etc. Analysis of Western blots suggests, however, that the normal presenilin proteins undergo proteolytic cleavage to yield an approximately 35 kDa N-terminal fragment and an approximately 18 kDa C-terminal fragment. In particular, Western blots bearing lysates from wild-type native human fibroblasts, human neocortical brain tissue from control subjects, and neocortical brain tissue from non-transgenic and PS1 transgenic mice using antibodies ("14.2") recognizing PS1-specific residues 1-25 at the N-terminus reveal the presence of a strong immunoreactive band of approximately 35 kDa and, after longer exposures, a weaker band of approximately 45 kDa which presumably represents the full-length PS1 protein. Antibodies ("520") directed at residues 304-318 at the apex of the TM6→7 loop of PS1, and antibodies ("4627") directed at residues 457-467 in the C-terminus of PS1, both recognize the same strong band of approximately 18 kDa. Antibodies 520 also recognize a weak band of 45 kDa

coincident with the PS1 band detected by 14.2. Sequencing of the major C-terminal fragment from PS1-transfected human embryonic kidney cells (HEK 293) showed that the principal endoproteolytic cleavage occurs near M298 in the proximal portion of the TM6→7 loop, possibly by enzymes other than the proteasome. These observations suggest that an endoproteolytic cleavage event occurs near the junction of exons 9 and 10 of PS1. Full length PS1 in these cells is quickly turned over ($t_{1/2}$ < 60 min.) by the proteasome.

To determine whether mutations in the presenilin proteins result in alterations of their proteolytic cleavage, Western blots containing lysates of fibroblast and neocortical brain homogenates from normal subjects and subjects carrying PS1 mutations were investigated with the PS1 specific antibody Ab 14.2. In fibroblasts, there were no obvious differences in the relative intensities of the protein bands when lysates from heterozygous carriers of the PS1 mutations were compared with normal homozygotes. In contrast, there appeared to be a difference between PS1 mutation carriers and normals in homogenates of temporal neocortex from AD affected heterozygous carriers of either the PS1 A246E or C410Y mutations (which are located in TM6 and TM7 respectively). In heterozygotes, a strongly immunoreactive band of approximately 45 kDa was detected which initially appeared to correspond to the full-length PS1 protein. Further analysis, however, revealed that this band represents an alternatively processed presenilin product. A similar band corresponding to this mutant processed PS1 was observed in neocortical homogenates from some sporadic late-onset AD patients. These data suggest that (1) some pathogenic PS1 mutations associated with early-onset AD alter the way in which the presenilins are processed through endoproteolytic and proteasome pathways and (2) the presenilin proteins, and changes in the processing of the presenilins in the brain, are also implicated in late-onset and sporadic AD.

2. Presenilin-Interacting Proteins

In order to identify proteins which may bind to or otherwise interact with the presenilins in vivo, a yeast two-hybrid system was used as described below

(Example 1). In particular, because mutations in the TM6→7 loop domains are known to be causative of AD, a yeast two-hybrid system was used to identify cellular proteins which may interact with normal and mutant presenilin TM6→7 loop domains. Yeast two-hybrid studies were also done with cDNAs corresponding to the C-terminal 18 kDa endoproteolytic cleavage fragment, and with cDNAs corresponding to the TM1→2 intraluminal loop domain, which is also the site of the FAD associated Y115H missense mutation. In brief, cDNA sequences encoding the TM6→7 loop (i.e., residues 266 to 409 of PS1) were ligated in-frame to the GAL4 DNA-binding domain in the pAS2-1 yeast expression plasmid vector (Clontech). This plasmid was then co-transformed into S. cerevisiae strain Y190 together with a library of human brain cDNAs ligated into the pACT2 yeast expression vector bearing the GAL4 activation domain (Clontech). After appropriate selection and re-screening, a number of clones were recovered and sequenced bearing human brain cDNAs encoding peptides which interacted with the normal presenilin TM6→7 domain. To determine whether these presenilin interactions would be modified by AD related mutations within the TM6→7 loop, the yeast two-hybrid system was again used with TM6→7 loop peptides containing the L286V, the L392V, and the exon 10 splicing mutants. When these mutant constructs were used as "bait" to re-screen the brain cDNA:GAL4 activation domain library, some but not all of the brain cDNA sequences which interacted with the normal presenilin were recovered. In addition, several new clones were identified which interacted with the mutant but not the normal presenilins. The clones corresponding to the PS-interacting proteins with the highest presenilin affinity are described in Example 1 and below.

PS-interacting proteins, particularly those which interact selectively with either the normal or mutant presenilins, provide new targets for the identification of useful pharmaceuticals, new targets for diagnostic tools in the identification of individuals at risk, new sequences for the production of transformed cell lines and transgenic animal models, and new bases for therapeutic intervention in Alzheimer's Disease. In particular, the onset of AD may be associated with aberrant interactions between mutant presenilin proteins and normal forms of PS-interacting proteins such

as those identified using the methods described herein. These changes may increase or decrease interactions present with normal PS1 or may cause interaction with a novel mutation-specific PS-interacting protein. In addition, however, aberrant interactions may result from normal presenilins binding to mutant forms of the PS-interacting proteins and, therefore, mutations in the PS-interacting proteins may also be causative of AD.

A. The S5a Subunit of the 26S Proteasome

Two overlapping clones have been identified as representing a portion of the human protein alternatively known as Antisecretory Factor ("ASF") or the Multiubiquitin chain-binding S5a subunit of the 26S proteasome ("S5a"). These clones, which together include residues 70-377 of S5a, were shown to interact with the normal presenilin TM6→7 loop domain but only weakly with two TM6→7 loop domain mutants tested (L286V, L392V). The PS1:S5a interaction was confirmed by co-immunoprecipitation studies, and immunocytochemical studies showed S5a and PS1 are expressed in contiguous intracellular compartments in brain cells typically affected by AD.

The interaction between PS1 and the proteasome could be relevant to the pathogenesis of Alzheimer's Disease (AD) through several possible mechanisms. First, most mammalian cells seem to maintain very low levels of the PS1 holoprotein. A notable exception to this are cells expressing the PS1 Δ290-319 splicing mutation, which results in a mutant PS1 holoprotein which is not endoproteolytically cleaved and which is, therefore, readily detectable. In the case of the Δ290-319 splicing mutation at least, the presence of the mutant PS1 holoprotein, or the absence or reduction in the 35 kDa N-terminal and 18 kDa C-terminal fragments, appears sufficient to cause AD. It is possible, therefore, that even very subtle changes in the turnover of the mutant PS1 holoprotein might have significant pathophysiological effects. Thus, mutations in either the presenilins or S5a which perturb the PS1:S5a interaction in the mammalian CNS may cause the presenilin holoprotein to be aberrantly processed and cause AD. Therefore, modulation of presenilin proteolytic pathways might be applied therapeutically to enhance removal of mutant holoprotein.

To assess a potential in vivo relationship between PS1 and the S5a subunit of the 26S proteasome, the effects of proteasome inhibitors on PS1 metabolism were investigated. Short term organotypic cultures of neonatal rat hippocampus and carcinoma of colon (CaCo2) cells (which express high levels of both PS1 and PS2) were administered either the specific, reversible proteasome inhibitor N-acetyl-leucinyl-leucinyl-norleucinyl-H (LLnL) (Rock et al., 1994), or the specific irreversible proteasome inhibitor lactacystin (Fenteany et al., 1995). Both agents caused an increase in the steady state levels of PS1 holoprotein. Both agents also prolonged the half-life of the PS1 holoprotein in pulse chase experiments in hippocampal slices from ~15 minutes to ~35 minutes. As noted above, the PS1 holoprotein appears to be rapidly turned over in normal cells. However, even after four hours of metabolic labelling, neither of the proteasome inhibitors affected the level of the 35 kDa N-terminal PS1 fragment, or resulted in the appearance of novel species. These studies imply that the majority of the PS1 holoprotein is catabolized directly via a rapid, proteasome dependent pathway in a manner similar to several other integral membrane proteins (e.g. Sec61 and CFTR). On the other hand, because the ~35 kDa and ~ 18 kDa terminal fragments are still produced in the presence of proteasome inhibitors, this endoproteolytic cleavage of PS1 is probably not mediated by the proteasome pathway. Therefore, it appears that at least two proteolytic pathways act upon the PS1 holoprotein.

An alternate possibility is that mutant PS1:S5a interactions may modify the function or the cellular regulation of S5a. To address this possibility, S5a levels were examined by Western blotting of lysates from postmortem temporal neocortex from non-AD neurologic controls (n = 8), sporadic AD (n = 8) and PS1-linked FAD (n = 4). In the majority of non-AD brains, polyclonal anti-S5a antibodies specifically detected an S5a species with Mr of ~ 50 kDa, which could be abolished by preabsorption of the antibody with recombinant His_6-S5a or with extracts of myc-S5a transfected cells. In a subset of these control cases an additional S5a reactive band was observed at ~34 kDa. In contrast, in tissue from all subjects with sporadic late onset AD, the predominant S5a reactive species was observed at ~ 40 kDa which was

not seen in control tissue. The origin, and the functional significance of this altered electrophoretic mobility is unclear but indicates that S5a processing is altered in AD brains, irrespective of whether the AD is presenilin-linked or sporadic.

Thus, the presenilin-proteasome interaction appears significant in several respects. First, the facts that the normal presenilin TM6→7 loop domain interacts with the S5a protein, that the mutant presenilin TM6→7 loop domains fail to interact (or interact very weakly) with the S5a protein, that presenilins bearing mutations in the TM6→7 loop domain appear to be differently cleaved and multiubiquitinated, that proteasomes are known to be involved in the cleavage and clearance of a variety of proteins (particularly multiubiquitinated proteins), that inhibition of proteasome activity inhibits cleavage of the presenilin holoproteins, and that S5a processing is altered in AD brains, all suggest (1) that the S5a subunit and the 26S proteasome are involved in the normal processing of the presenilins and that mutations which disrupt this normal interaction may be responsible for the abnormal processing observed in TM6→7 loop domain mutants; or (2) that the presenilin-proteasome interaction may modulate the activity of PS1, S5a, or both, with or without involving proteasome-mediated presenilin processing; or (3) that modulation of the normal quality control function of proteasome-mediated degradation of misfolded or mutant membrane proteins trafficking through the ER and Golgi (such as APP, Notch, or Prion proteins), and of misfolded, mutant, or ubiquitinated cytoplasmic proteins (including structural proteins such as tau, and short lived, proteasome processed signaling molecules such as NFkB). Thus, defective proteasome function might selectively cause these proteins (especially βAPP, tau, Prion) to be aberrantly metabolized. The latter would lead to the accumulation of neurotoxic, amyloidogenic protease-resistant derivatives such as Aβ and PrPsc, the accumulation of neurofibrillary tangles, and defective intracellular signaling functions. In support of these hypotheses, it should be noted that failure to clear hyperubiquitinated phosphorylated tau and other microtubule associated proteins is a prominent feature of Alzheimer's Disease (Kosik and Greenberg, 1994), suggesting a possible link between TM6→7 loop domain mutants, presenilin-proteasome interactions, tau-proteasome interactions, and the neurofibrillary tangles

of tau protein in AD brains. Finally, proteasomes are known to be capable of degrading APP and of binding the Aβ peptides which are associated with Alzheimer's Disease, suggesting a possible link between TM6→7 loop domain mutants, presenilin-proteasome interactions, APP-proteasome interactions, and the amyloid plaques characteristic of AD brains. Furthermore, administration of proteasome inhibitors such as LLnL and Lactacystin cause severe disturbances in βAPP metabolism with increases in intracellular immature N-glycosylated βAPP, and the secretion of much larger amounts of $A\beta_{42}$ isoforms into the media (Klafki, et al., 1996).

Therefore, presenilin processing and the presenilin-proteasome interaction are clear targets for the diagnosis as well as therapeutic intervention in AD. Thus, as described below, assays may now be provided for drugs which affect the proteasome-mediated cleavage of the presenilins, which affect the alternative endoproteolytic cleavage and ubiquitination of the mutant presenilins, or which otherwise affect the processing and trafficking of the presenilins or the S5a subunit of the proteasome. In addition, as mutations in the 26S proteasome which disrupt the normal processing of the presenilins are likely to be causative of Alzheimer's Disease, additional diagnostic assays are provided for detecting mutations in the S5a or other subunits of the proteasome. Finally, additional transformed cell lines and transgenic models may now be provided which have been altered by the introduction of a normal or mutant sequence encoding at least a functional domain of the proteasome. The appearance of abnormal electrophoretic forms of S5a (and/or other proteasome subunits) in biologic tissues and fluids can be used as a clinical test for diagnosis and monitoring of disease activity in subjects with sporadic forms of AD.

B. GT24: A Protein with "Armadillo" Repeats

Another PS-interacting protein, designated GT24, was identified from several over-lapping clones obtained using a $PS1_{266-409}$ domain as bait in the yeast two-hybrid system and a human adult brain cDNA library. Six longer GT24 clones of ~3.8 kb in size were subsequently obtained by screening of conventional cDNA libraries. The open reading frame within the longest GT24 clone obtained to date

(Accession number U81004) suggests that GT24 is a protein of at least 1040 amino acids with a unique N-terminus, and considerable homology to several armadillo (arm) repeat proteins at its C-terminus. Thus, for example, residues 440-862 of GT24 (numbering from Accession number U81004) have 32-56% identity ($p=1.2e^{-133}$) to residues 440-854 of murine p120 protein (Accession number Z17804), and residues 367-815 of GT24 have 26-42% identity (p=0.0017) to residues 245-465 of the D. melanogaster armadillo segment polarity protein (Accession number P18824). The GT24 gene maps to chromosome 5p15 near the anonymous microsatellite marker D5S748 and the Cri-du-Chat syndrome locus.

Hybridization of unique 5' sequences of GT24 to Northern blots reveals that the GT24 gene is expressed as a range of transcripts varying in size between ~3.9 and 5.0 kb in several regions of human brain, and in several non-neurologic tissues such as heart. In addition, in situ hybridization studies using a 289 bp single copy fragment from the 5' end of GT24 in four month old murine brain reveal GT24 transcription closely parallels that of PS1, with robust expression in dentate and hippocampal neurons, in scattered neocortical neurons, and in cerebellar Purkinje cells. In day E13 murine embryos, GT24 is widely expressed at low levels, but is expressed at somewhat higher levels in somites and in the neural tube. A physiological in vivo interaction between GT24 and PS1 is supported by co-immunoprecipitation studies in HEK293 cells transiently transfected with a wild type human PS1 cDNA, a c-myc-tagged cDNA encoding residues 484-1040 of GT24 (including the C-terminal arm repeats), or both cDNAs. Cell lysates were immunoprecipitated with anti-PS1 antibodies and then investigated for the presence of the myc-GT24 protein by immuno-blotting. In PS1/myc-GT24 double transfected cells, the immunoprecipitates contained a robust anti-myc reactive band of Mr ~60 kDa, which co-migrated with a myc-GT24 control. In cells transfected with myc-GT24 only, a very weak band was detected after long exposures, presumably reflecting interaction of the myc-GT24 with low levels of endogenous PS1. No myc-reactive bands were detected in cells transfected with PS1 alone, or in any of the transfected cells immunoprecipitated with pre-immune serum. Taken together, these

observations strongly suggest that the observed PS1:GT24 interaction is physiologically relevant.

To explore whether mutations in the TM6-TM7 loop of PS1 might influence the PS1:GT24 interaction, we employed quantitative liquid β-galactosidase assays to directly compare the yeast-two-hybrid interaction of the C-terminal residues 499-1040 of GT24 with wildtype and mutant $PS1_{266-409}$. These studies revealed that the interaction of $GT24_{499-1040}$ with a L286V mutant PS1 domain was not significantly different from the interaction with the corresponding wild type PS1 domain. In contrast, there was a significant reduction in the $GT24_{499-1040}$ interaction with the L392V mutant PS1 construct. The absence of an effect of the L286V mutation, and the presence of an effect with the L392V mutation, may suggest that some mutations may effect PS1:GT24 binding, while others may modulate the PS1 response to GT24 binding.

The PS1:GT24 interaction could support several functions. The arm repeat motif of GT24 has been detected in several proteins with diverse functions including β-catenin and its invertebrate homologue armadillo, plakoglobin, p120, the adenomatous polyposis coli (APC) gene, suppressor of RNA polymerase 1 in yeast (SRP1), and smGDS. For example, β-catenin, p120 and plakoglobin play an essential role in intercellular adhesion. β-catenin/armadillo is involved in transduction of wingless/Wnt signals during cell fate specification, and β-catenin and p120 may play a role in other receptor mediated signal transduction events including responses to trophic factors such as PDGF, EGF, CSF-1 and NGF.

If the PS1:GT24 interaction is part of intercellular signaling pathways for trophic factors, or is involved in cell-cell adherence, disruption of the interaction may be involved in the neurodegenerative processes in PS-linked FAD brains, and in the increased sensitivity of PS1 or PS2 transfected cells to apoptosis (Wolozin et al., 1996). It is of note that at least one arm protein, smGDS, stimulates GDP/GTP exchange on intracellular G-proteins (Kikuchi et al. 1992; Borguski et al., 1993), and that mutant forms of both βAPP and PS2 are thought to activate programmed cell

death pathways through mechanisms involving heterotrimeric GTP/GDP proteins (Wolozin et al, 1996; Okamoto, et al., 1995; Yamatsuji, et al, 1996).

The interaction between PS1 and GT24 may also be involved in some of the developmental phenotypes associated with homozygous PS1 knockouts in mice such as failed somitogenesis of the caudal embryo, short tail, and fatal cerebral hemorrhage at around day E13.5 (Wong et al., 1996). The resemblance of these skeletal phenotypes to those associated with null mutations in PAX1 and Notch, and the apparent suppressor effect of mutations in sel12 on Notch/lin12 mediated signaling in C. elegans suggest that the PS proteins function in the Notch signaling pathway. In addition, mice homozygous for a knockout of the Wnt-3a gene (Takada et al., 1994), and murine homozygotes for a spontaneous mutation, "vestigial tail" or vt, in the Wnt-3a gene (Greco et al., 1996), have skeletal phenotypes of defective caudal somite and tail bud formation. The Wnt-3a knockouts are embryonic lethal by day 12.5. These phenotypes are similar to those of homozygous knockouts of the murine PS1 gene (Wong et al., 1996). The observation that GT24 binds to PS1, is expressed in embryonic somites, and contains the armadillo repeat motif of other proteins used in the downstream signaling in the Wingless/Wnt pathway suggests that PS1 is a downstream element in the GT24-Wingless/Wnt pathway. This can be exploited to create a bioassay for drugs affecting the GT24-PS1 interaction directly, or affecting upstream or downstream components of that interaction pathway, and can therefore be used to monitor the effects of presenilin mutations. For example, cells transfected with normal or mutant presenilins may be exposed to soluble Wnt-3a protein (or other Wnt proteins such as Wnt-1) and assayed for changes which are specific to the Wingless/Wnt signaling pathway, or for any of the other changes described herein for cell assays (e.g., intracellular ion levels, Aβ processing, apoptosis, etc.).

Thus, the GT24 protein also presents new targets for diagnosis as well as therapeutic intervention in AD. For example, as mutations in the GT24 protein may also be causative of Alzheimer's Disease, additional diagnostic assays are provided for detecting mutations in these sequences. Similarly, additional transformed cell lines

and transgenic models may now be provided which have been altered by introduction of a normal or mutant nucleic acid encoding at least a functional domain of the GT24 protein, and particularly the functional domains (e.g., residues 70-377) which interact with the presenilins. Such transformed cells and transgenics will have utility in assays for compounds which modulate the presenilin-GT24 interactions.

C. p0071: A Protein with "Armadillo" Repeats

Another independent clone isolated in the initial screening with the wild type $PS1_{266-409}$ "bait" also encodes a peptide with C-terminal arm repeats (clone Y2H25, Accession number U81005). A longer cDNA sequence corresponding to the Y2H25 clone has been deposited with GenBank as human protein p0071 (Accession number X81889) and is reproduced herein as SEQ ID NO: 5. Clone Y2H25 corresponds essentially to nucleotide positions 1682-1994 of SEQ ID NO: 5. Comparison of the predicted sequence of the Y2H25/p0071 ORF with that of GT24 confirms that they are related proteins with 47% overall amino acid sequence identity, and with 70% identity between residues 346-862 of GT24 and residues 509-1022 of p0071. This suggests that PS1 interacts with a novel class of arm repeat containing proteins. The broad ~4.5 kb hybridization signal obtained on Northern blots with the unique 5' end of GT24 could reflect either alternative splicing/polyadenylation of GT24 or, less likely, the existence of additional members of this family with higher degrees of N-terminal homology to GT24 than p0071. Cells transformed with these sequences, or transgenic animals including these sequences, will have additional utility as animal models of AD and for use in screening for compounds which modulate the action of normal and mutant presenilins.

D. Rab 11

One clone (Y2H9), disclosed herein as SEQ ID NO: 5, was identified as interacting with the normal PS1 TM6→7 loop domain and appears to correspond to a known gene, Rab11, available through Accession numbers X56740 and X53143. Rab11 is believed to be involved in protein/vesicle trafficking in the ER/Golgi. Note the possible relationship to processing of membrane proteins such as βAPP and Notch

with resultant overproduction of toxic Aß peptides (especially neurotoxic $A\beta_{1-42(43)}$ isoforms) (Scheuner, et al, 1995).

E. Retinoid X Receptor-β

One clone (Y2H23b), disclosed herein as SEQ ID NO: 6, was identified as interacting with the normal PS1 TM6→7 loop domain and appears to correspond to a known gene, known variously as the retinoid X receptor-β, nuclear receptor co-regulator, or MHC Class I regulatory element, and is available through Accession numbers M84820, X63522 and M81766. This gene is believed to be involved in intercellular signaling, suggesting a possible relationship to the intercellular signaling function mediated by C. elegans sel12 and Notch/lin-12 (transcription activator).

F. Cytoplasmic Chaperonin

One clone (Y2H27), disclosed herein as SEQ ID NO: 8, was identified as interacting with the normal PS1 TM6→7 loop domain and appears to correspond to a known gene, a cytoplasmic chaperonin containing TCP-1, available through Accession numbers U17104 and X74801.

G. Clone Y2H35

One clone (Y2H35), disclosed herein as SEQ ID NO: 7, was identified as interacting with the normal PS1 TM6→7 loop domain and appears to correspond to a sequence that codes for a protein of unknown function, available through Accession number R12984, but which displays evolutionary conservation in yeast sequences.

H. Clone Y2H171

One clone (Y2H171), disclosed herein as SEQ ID NO: 9, was identified as interacting with the normal PS1 TM6→7 loop domain and appears to correspond to a known expressed repeat sequence available through Accession number D55326.

I. Clone Y2H41

One clone (Y2H41) was identified which reacts strongly with the TM6→7 loop domains of both PS1 and PS2 as well as the mutant loop domains of PS1. The sequence, disclosed as SEQ ID NO: 10, shows strong homology to an EST of unknown function (Accession number T64843).

III. Preferred Embodiments

Based, in part, upon the discoveries disclosed and described herein, the following preferred embodiments of the present invention are provided.

1. Isolated Nucleic Acids

In one series of embodiments, the present invention provides isolated nucleic acids corresponding to, or relating to, the nucleic acid sequences disclosed herein, which encode at least the PS-interacting domain of a PS-interacting protein. As described more fully below, the disclosed and enabled sequences include normal sequences from humans and other mammalian species, mutant sequences from humans and other mammalian species, homologous sequences from non-mammalian species such as Drosophila and C. elegans, subsets of these sequences useful as probes and PCR primers, subsets of these sequences encoding fragments of the PS-interacting proteins or corresponding to particular structural domains or polymorphic regions, complementary or antisense sequences corresponding to fragments of the PS-interacting protein genes, sequences in which the PS-interacting protein coding regions have been operably joined to exogenous regulatory regions, and sequences encoding fusion proteins in which portions of the PS-interacting proteins are fused to other proteins useful as markers of expression, as "tags" for purification, or in screens and assays for other proteins which interact with the PS-interacting proteins.

Thus, in a first series of embodiments, isolated nucleic acid sequences are provided which encode at least a PS-interacting domain of a normal or mutant version of a PS-interacting protein. Examples of such nucleic acid sequences are disclosed herein as SEQ ID NOs: 1, 3, and 5-10. In addition, given the sequences of the PS-interacting domains of the PS-interacting proteins disclosed herein, one of ordinary skill in the art is clearly enabled to obtain the entire genomic or cDNA sequence encoding the entire PS-interacting proteins. Thus, for example, based upon the initial clone of the GT24 protein obtained using the yeast two-hybrid system (Example 1), the larger GT24 clone disclosed as SEQ ID NO: 3 was obtained by standard methods

known in the art. Complete cDNA or genomic clones of each of the genes encoding the disclosed sequences may be similarly obtained by one of ordinary skill in the art. Therefore, the present invention provides complete genomic sequences as well as cDNA sequences corresponding to the PS-interacting protein genes of the invention. Alternatively, the nucleic acids of the invention may comprise recombinant genes or "minigenes" in which all or some introns of the PS-interacting protein genes have been removed, or in which various combinations of introns and exons and local cis-acting regulatory elements have been engineered in propagation or expression constructs or vectors. For purposes of reducing the size of a recombinant PS-interacting protein gene, a cDNA gene may be employed, or various combinations of introns and untranslated exons may be removed from a DNA construct. These and many variations on these embodiments are now enabled by the identification and description of the PS-interacting proteins provided herein.

In addition to the disclosed PS-interacting protein and gene sequences, one of ordinary skill in the art is now enabled to identify and isolate nucleic acids representing PS-interacting genes or cDNAs which are allelic to the disclosed sequences or which are heterospecific homologues. Thus, the present invention provides isolated nucleic acids corresponding to these alleles and homologues, as well as the various above-described recombinant constructs derived from these sequences, by means which are well known in the art. Briefly, one of ordinary skill in the art may now screen preparations of genomic or cDNA, including samples prepared from individual organisms (e.g., human AD patients or their family members) as well as bacterial, viral, yeast or other libraries of genomic or cDNA, using probes or PCR primers to identify allelic or homologous sequences. Because it is desirable to identify mutations in the PS-interacting proteins which may contribute to the development of AD or other disorders, because it is desirable to identify polymorphisms in the PS-interacting proteins which are not pathogenic, and because it is also desirable to create a variety of animal models which may be used to study AD and screen for potential therapeutics, it is particularly contemplated that additional PS-interacting protein sequences will be isolated from other preparations or libraries

of human nucleic acids and from preparations or libraries from animals including rats, mice, hamsters, guinea pigs, rabbits, dogs, cats, goats, sheep, pigs, and non-human primates. Furthermore, PS-interacting protein homologues from yeast or invertebrate species, including C. elegans and other nematodes, as well as Drosophila and other insects, may have particular utility for drug screening.

Standard hybridization screening or PCR techniques may be employed (as used, for example, in the identification of the mPS1 gene disclosed in PCT Publication WO96/34099) to identify and/or isolate such allelic and homologous sequences using relatively short PS-interacting protein gene sequences. The sequences may include 8 or fewer nucleotides depending upon the nature of the target sequences, the method employed, and the specificity required. Future technological developments may allow the advantageous use of even shorter sequences. With current technology, sequences of 9-50 nucleotides, and preferably about 18-24 are preferred. These sequences may be chosen from those disclosed herein, or may be derived from other allelic or heterospecific homologues enabled herein. When probing mRNA or screening cDNA libraries, probes and primers from coding sequences (rather than introns) are preferably employed, and sequences which are omitted in alternative splice variants typically are avoided unless it is specifically desired to identify those variants. Allelic variants of the PS-interacting protein genes may be expected to hybridize to the disclosed sequences under stringent hybridization conditions, as defined herein, whereas lower stringency may be employed to identify heterospecific homologues.

In another series of embodiments, the present invention provides for isolated nucleic acids which include subsets of the PS-interacting protein sequences or their complements. As noted above, such sequences will have utility as probes and PCR primers in the identification and isolation of allelic and homologous variants of the PS-interacting protein genes. Subsequences corresponding to polymorphic regions of the PS-interacting proteins, will also have particular utility in screening and/or genotyping individuals for diagnostic purposes, as described below. In addition, and also as described below, such subsets will have utility for encoding (1)

fragments of the PS-interacting proteins for inclusion in fusion proteins, (2) fragments which comprise functional domains of the PS-interacting proteins for use in binding studies, (3) fragments of the PS-interacting proteins which may be used as immunogens to raise antibodies against the PS-interacting proteins, and (4) fragments of the PS-interacting proteins which may act as competitive inhibitors or as mimetics of the PS-interacting proteins to inhibit or mimic their physiological functions. Finally, such subsets may encode or represent complementary or antisense sequences which can hybridize to the PS-interacting protein genes or PS-interacting protein mRNA transcripts under physiological conditions to inhibit the transcription or translation of those sequences. Therefore, depending upon the intended use, the present invention provides nucleic acid subsequences of the PS-interacting protein genes which may have lengths varying from 8-10 nucleotides (e.g., for use as PCR primers) to nearly the full size of the PS-interacting protein genomic or cDNAs. Thus, the present invention provides isolated nucleic acids comprising sequences corresponding to at least 8-10, preferably 15, and more preferably at least 20 consecutive nucleotides of the PS-interacting protein genes, as disclosed or otherwise enabled herein, or to their complements. As noted above, however, shorter sequences may be useful with different technologies.

In another series of embodiments, the present invention provides nucleic acids in which the coding sequences for the PS-interacting proteins, with or without introns or recombinantly engineered as described above, are operably joined to endogenous or exogenous 5' and/or 3' regulatory regions. Using the present disclosure and standard genetic techniques (e.g., PCR extensions, targeting gene walking), one of ordinary skill in the art is now enabled to clone the 5' and/or 3' endogenous regulatory regions of any of the disclosed PS-interacting protein genes. Similarly, allelic variants of these endogenous regulatory regions, as well as endogenous regulatory regions from other mammalian homologues, are similarly enabled without undue experimentation. Alternatively, exogenous regulatory regions (i.e., regulatory regions from a different conspecific gene or a heterospecific regulatory region) may be operably joined to the PS-interacting protein coding sequences in order to drive

expression. Appropriate 5' regulatory regions will include promoter elements and may also include additional elements such as operator or enhancer sequences, ribosome binding sequences, RNA capping sequences, and the like. The regulatory region may be selected from sequences that control the expression of genes of prokaryotic or eukaryotic cells, their viruses, and combinations thereof. Such regulatory regions include, but are not limited to, the lac system, the trp system, the tac system, and the trc system; major operator and promoter regions of phage λ; the control region of the fd coat protein; early and late promoters of SV40; promoters derived from polyoma, adenovirus, retrovirus, baculovirus, and simian virus; 3-phosphoglycerate kinase promoter; yeast acid phosphatase promoters; yeast alpha-mating factors; promoter elements of other eukaryotic genes expressed in neurons or other cell types; and combinations thereof. In particular, regulatory elements may be chosen which are inducible or repressible (e.g., the β-galactosidase promoter) to allow for controlled and/or manipulable expression of the PS-interacting protein genes in cells transformed with these nucleic acids. Alternatively, the PS-interacting protein coding regions may be operably joined with regulatory elements which provide for tissue specific expression in multicellular organisms. Such constructs are particularly useful for the production of transgenic organisms to cause expression of the PS-interacting protein genes only in appropriate tissues. The choice of appropriate regulatory regions is within the ability and discretion of one of ordinary skill in the art and the recombinant use of many such regulatory regions is now established in the art.

In another series of embodiments, the present invention provides for isolated nucleic acids encoding all or a portion of the PS-interacting proteins in the form of a fusion protein. In these embodiments, a nucleic acid regulatory region (endogenous or exogenous) is operably joined to a first coding region which is covalently joined in-frame to a second coding region. The second coding region optionally may be covalently joined to one or more additional coding regions and the last coding region is joined to a termination codon and, optionally, appropriate 3' regulatory regions (e.g., polyadenylation signals). The PS-interacting protein sequences of the fusion protein may represent the first, second, or any additional

coding regions. The PS-interacting protein sequences may be conserved or non-conserved domains and can be placed in any coding region of the fusion. The non-PS-interacting protein sequences of the fusion may be chosen according to the needs and discretion of the practitioner and are not limited by the present invention. Useful non-PS-interacting protein sequences include, for example, short sequence "tags" such as antigenic determinants or poly-His tags which may be used to aid in the identification or purification of the resultant fusion protein. Alternatively, the non-PS-interacting protein coding region may encode a large protein or protein fragment, such as an enzyme or binding protein which also may assist in the identification and purification of the protein, or which may be useful in an assay such as those described below. Particularly contemplated fusion proteins include poly-His and GST (glutathione S-transferase) fusions which are useful in isolating and purifying the presenilins-interacting proteins, and the yeast two hybrid fusions, described below, which are useful in assays to identify other proteins which bind to or interact with the PS-interacting proteins.

In another series of embodiments, the present invention provides isolated nucleic acids in the form of recombinant DNA constructs in which a marker or reporter gene (e.g., β-galactosidase, luciferase) is operably joined to the 5' regulatory region of a PS-interacting protein gene such that expression of the marker gene is under the control of those regulatory sequences. Using the PS-interacting protein regulatory regions enabled herein, including regulatory regions from human and other mammalian species, one of ordinary skill in the art is now enabled to produce such constructs. As discussed more fully below, such isolated nucleic acids may be used to produce cells, cell lines or transgenic animals which are useful in the identification of compounds which can, directly or indirectly, differentially affect the expression of the PS-interacting proteins.

Finally, the isolated nucleic acids of the present invention include any of the above described sequences when included in vectors. Appropriate vectors include cloning vectors and expression vectors of all types, including plasmids, phagemids, cosmids, episomes, and the like, as well as integration vectors. The vectors may also

include various marker genes (e.g., antibiotic resistance or susceptibility genes) which are useful in identifying cells successfully transformed therewith. In addition, the vectors may include regulatory sequences to which the nucleic acids of the invention are operably joined, and/or may also include coding regions such that the nucleic acids of the invention, when appropriately ligated into the vector, are expressed as fusion proteins. Such vectors may also include vectors for use in yeast "two hybrid," baculovirus, and phage-display systems. The vectors may be chosen to be useful for prokaryotic, eukaryotic or viral expression, as needed or desired for the particular application. For example, vaccinia virus vectors or simian virus vectors with the SV40 promoter (e.g., pSV2), or Herpes simplex virus or adeno-associated virus may be useful for transfection of mammalian cells including neurons in culture or in vivo, and the baculovirus vectors may be used in transfecting insect cells (e.g., butterfly cells). A great variety of different vectors are now commercially available and otherwise known in the art, and the choice of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

2. Substantially Pure Proteins

The present invention provides for substantially pure preparations of the PS-interacting proteins, fragments of the PS-interacting proteins, and fusion proteins including the PS-interacting proteins or fragments thereof. The proteins, fragments and fusions have utility, as described herein, in the generation of antibodies to normal and mutant PS-interacting proteins, in the identification of proteins (aside from the presenilins) which bind to the PS-interacting proteins, and in diagnostic and therapeutic methods. Therefore, depending upon the intended use, the present invention provides substantially pure proteins or peptides comprising amino acid sequences which are subsequences of the complete PS-interacting proteins and which may have lengths varying from 4-10 amino acids (e.g., for use as immunogens), or 10-100 amino acids (e.g., for use in binding assays), to the complete PS-interacting proteins. Thus, the present invention provides substantially pure proteins or peptides comprising sequences corresponding to at least 4-5, preferably 6-10, and more

preferably at least 50 or 100 consecutive amino acids of the PS-interacting proteins, as disclosed or otherwise enabled herein.

The proteins or peptides of the invention may be isolated and purified by any of a variety of methods selected on the basis of the properties revealed by their protein sequences. For example, the PS-interacting proteins may be isolated from cells in which the PS-interacting protein is normally highly expressed. Alternatively the PS-interacting protein, fusion protein, or fragment thereof, may be purified from cells transformed or transfected with expression vectors (e.g., baculovirus systems such as the pPbac and pMbac vectors (Stratagene, La Jolla, CA); yeast expression systems such as the pYESHIS Xpress vectors (Invitrogen, San Diego, CA); eukaryotic expression systems such as pcDNA3 (Invitrogen, San Diego, CA) which has constant constitutive expression, or LacSwitch (Stratagene, La Jolla, CA) which is inducible; or prokaryotic expression vectors such as pKK233-3 (Clontech, Palo Alto, CA). In the event that the protein or fragment integrates into the endoplasmic reticulum or plasma membrane of the recombinant cells (e.g., eukaryotic cells), the protein may be purified from the membrane fraction. Alternatively, if the protein aggregates in inclusion bodies within the recombinant cells (e.g., prokaryotic cells), the protein may be purified from whole lysed cells or from solubilized inclusion bodies.

Purification can be achieved using standard protein purification procedures including, but not limited to, gel-filtration chromatography, ion-exchange chromatography, high-performance liquid chromatography (RP-HPLC, ion-exchange HPLC, size-exclusion HPLC, high-performance chromatofocusing chromatography, hydrophobic interaction chromatography, immunoprecipitation, or immunoaffinity purification. Gel electrophoresis (e.g., PAGE, SDS-PAGE) can also be used to isolate a protein or peptide based on its molecular weight, charge properties and hydrophobicity.

A PS-interacting protein, or a fragment thereof, may also be conveniently purified by creating a fusion protein including the desired PS-interacting protein sequence fused to another peptide such as an antigenic determinant or poly-His tag (e.g., QIAexpress vectors, QIAGEN Corp., Chatsworth, CA), or a larger protein (e.g.,

GST using the pGEX-27 vector (Amrad, USA) or green fluorescent protein using the Green Lantern vector (GIBCO/BRL. Gaithersburg, MD). The fusion protein may be expressed and recovered from prokaryotic or eukaryotic cells and purified by any standard method based upon the fusion vector sequence. For example, the fusion protein may be purified by immunoaffinity or immunoprecipitation with an antibody to the non-PS-interacting protein portion of the fusion or, in the case of a poly-His tag, by affinity binding to a nickel column. The desired PS-interacting protein or fragment may then be further purified from the fusion protein by enzymatic cleavage of the fusion protein. Methods for preparing and using such fusion constructs for the purification of proteins are well known in the art and several kits are commercially available for this purpose. In light of the present disclosure, one is now enabled to employ such fusion constructs with the PS-interacting proteins.

3. Antibodies to the PS-interacting Proteins

The present invention also provides antibodies, and methods of making antibodies, which selectively bind to the PS-interacting proteins or fragments thereof. Of particular importance, by identifying the PS-interacting domains of the PS-interacting proteins, and methods of identifying mutant forms of the PS-interacting proteins associated with Alzheimer's Disease, the present invention provides antibodies, and methods of making antibodies, which will selectively bind to and, thereby, identify and/or distinguish normal and mutant (i.e., pathogenic) forms of the PS-interacting proteins. The antibodies of the invention have utility as laboratory reagents for, inter alia, immunoaffinity purification of the PS-interacting proteins, Western blotting to identify cells or tissues expressing the PS-interacting proteins, and immunocytochemistry or immunofluorescence techniques to establish the subcellular location of the proteins. In addition, as described below, the antibodies of the invention may be used as diagnostics tools to identify carriers of AD-related PS-interacting protein alleles, or as therapeutic tools to selectively bind and inhibit pathogenic forms of the PS-interacting proteins in vivo.

The antibodies of the invention may be generated using the entire PS-interacting proteins of the invention, or using any PS-interacting protein epitope which is characteristic of that protein and which substantially distinguishes it from other host proteins. Any method of choosing antigenic determinants known in the art may, of course, be employed. Such epitopes may be identified by comparing sequences of, for example, 4-10 amino acid residues from a PS-interacting protein sequence to computer databases of protein sequences from the relevant host. In addition, larger fragments (e.g., 8-20 or, preferably, 9-15 residues) including one or more potential epitopes may also be employed. Antibodies to the PS-interacting domains (identified by the yeast two-hybrid assays described below) are expected to have the greatest utility both diagnostically and therapeutically. On the other hand, antibodies against highly conserved domains are expected to have the greatest utility for purification or identification of PS-interacting proteins.

PS-interacting protein immunogen preparations may be produced from crude extracts (e.g., lysates or membrane fractions of cells highly expressing the proteins), from proteins or peptides substantially purified from cells which naturally or recombinantly express them or, for short immunogens, by chemical peptide synthesis. The immunogens may also be in the form of a fusion protein in which the non-PS-interacting protein region is chosen for its adjuvant properties. As used herein, a PS-interacting protein immunogen shall be defined as a preparation including a peptide comprising at least 4-8, and preferably at least 9-15 consecutive amino acid residues of a PS-interacting proteins, as disclosed or otherwise enabled herein. Sequences of fewer residues may, of course, also have utility depending upon the intended use and future technological developments. Therefore, any PS-interacting protein derived sequences which are employed to generate antibodies to the PS-interacting proteins should be regarded as PS-interacting protein immunogens.

The antibodies of the invention may be polyclonal or monoclonal, or may be antibody fragments, including Fab fragments, $F(ab')_2$, and single chain antibody fragments. In addition, after identifying useful antibodies by the method of the invention, recombinant antibodies may be generated, including any of the antibody

fragments listed above, as well as humanized antibodies based upon non-human antibodies to the PS-interacting proteins. In light of the present disclosure, as well as the characterization of other PS-interacting proteins enabled herein, one of ordinary skill in the art may produce the above-described antibodies by any of a variety of standard means well known in the art. For an overview of antibody techniques, see Antibody Engineering: A Practical Guide, Borrebaek, ed., W.H. Freeman & Company, NY (1992), or Antibody Engineering, 2nd Ed., Borrebaek, ed., Oxford University Press, Oxford (1995).

As a general matter, polyclonal antibodies may be generated by first immunizing a mouse, rabbit, goat or other suitable animal with the PS-interacting protein immunogen in a suitable carrier. To increase the immunogenicity of the preparation, the immunogen may be coupled to a carrier protein or mixed with an adjuvant (e.g., Freund's adjuvant). Booster injections, although not necessary are recommended. After an appropriate period to allow for the development of a humoral response, preferably several weeks, the animals may be bled and the sera may be purified to isolate the immunoglobulin component.

Similarly, as a general matter, monoclonal anti-PS-interacting protein antibodies may be produced by first injecting a mouse, rabbit, goat or other suitable animal with a PS-interacting protein immunogen in a suitable carrier. As above, carrier proteins or adjuvants may be utilized and booster injections (e.g., bi- or tri-weekly over 8-10 weeks) are recommended. After allowing for development of a humoral response, the animals are sacrificed and their spleens are removed and resuspended in, for example, phosphate buffered saline (PBS). The spleen cells serve as a source of lymphocytes, some of which are producing antibody of the appropriate specificity. These cells are then fused with an immortalized cell line (e.g., myeloma), and the products of the fusion are plated into a number of tissue culture wells in the presence of a selective agent such as HAT. The wells are serially screened and replated, each time selecting cells making useful antibody. Typically, several screening and replating procedures are carried out until over 90% of the wells contain single clones which are positive for antibody production. Monoclonal antibodies

produced by such clones may be purified by standard methods such as affinity chromatography using Protein A Sepharose, by ion-exchange chromatography, or by variations and combinations of these techniques.

The antibodies of the invention may be labelled or conjugated with other compounds or materials for diagnostic and/or therapeutic uses. For example, they may be coupled to radionuclides, fluorescent compounds, or enzymes for imaging or therapy, or to liposomes for the targeting of compounds contained in the liposomes to a specific tissue location.

4. Transformed Cell Lines

The present invention also provides for cells or cell lines, both prokaryotic and eukaryotic, which have been transformed or transfected with the nucleic acids of the present invention so as to cause clonal propagation of those nucleic acids and/or expression of the proteins or peptides encoded thereby. Such cells or cell lines will have utility both in the propagation and production of the nucleic acids and proteins of the present invention but also, as further described herein, as model systems for diagnostic and therapeutic assays. In particular, it is expected that cells co-transformed with PS-interacting protein sequences as well as presenilin sequences will have improved utility as models of the biochemical pathways which may be affected in AD. For example, cells co-transformed with the interacting domains of PS-interacting sequences and presenilins in yeast two-hybrid fusion constructs, will have utility in screening for compounds which either enhance or inhibit interactions between these domains. Similarly, for cells transformed with a heterospecific presenilin, co-transformation with a similarly heterospecific PS-interacting protein, or co-transformation and homologous recombination to introduce a similarly heterospecific PS-interacting domain of a PS-interacting protein (e.g., "humanizing" a non-human endogenous PS-interacting protein), will result in a better model system for studying the interactions of the presenilins and the PS-interacting proteins. Cells transformed with only PS-interacting sequences will, of course, have utility of their

own for studying the role of these proteins in the etiology of AD, and also as precursors for presenilin co-transformed cells.

As used herein, the term "transformed cell" is intended to embrace any cell, or the descendant of any cell, into which has been introduced any of the nucleic acids of the invention, whether by transformation, transfection, infection, or other means. Methods of producing appropriate vectors, transforming cells with those vectors, and identifying transformants are well known in the art and are only briefly reviewed here (see, for example, Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

Prokaryotic cells useful for producing the transformed cells of the invention include members of the bacterial genera Escherichia (e.g., E. coli), Pseudomonas (e.g., P. aeruginosa), and Bacillus (e.g., B. subtillus, B. stearothermophilus), as well as many others well known and frequently used in the art. Prokaryotic cells are particularly useful for the production of large quantities of the proteins or peptides of the invention (e.g., normal or mutant PS-interacting proteins, fragments of the PS-interacting proteins, fusion proteins of the PS-interacting proteins). Bacterial cells (e.g., E. coli) may be used with a variety of expression vector systems including, for example, plasmids with the T7 RNA polymerase/promoter system, bacteriophage λ regulatory sequences, or M13 Phage mGPI-2. Bacterial hosts may also be transformed with fusion protein vectors which create, for example, lacZ, trpE, maltose-binding protein, poly-His tags, or glutathione-S-transferase fusion proteins. All of these, as well as many other prokaryotic expression systems, are well known in the art and widely available commercially (e.g., pGEX-27 (Amrad, USA) for GST fusions).

Eukaryotic cells and cell lines useful for producing the transformed cells of the invention include mammalian cells and cell lines (e.g., PC12, COS, CHO, fibroblasts, myelomas, neuroblastomas, hybridomas, human embryonic kidney 293, oocytes, embryonic stem cells), insect cells lines (e.g., using baculovirus vectors such as pPbac or pMbac (Stratagene, La Jolla, CA)), yeast (e.g., using yeast expression

vectors such as pYESHIS (Invitrogen, CA)), and fungi. Eukaryotic cells are particularly useful for embodiments in which it is necessary that the PS-interacting proteins, or functional fragments thereof, perform the functions and/or undergo the intracellular interactions associated with either the normal or mutant proteins. Thus, for example, transformed eukaryotic cells are preferred for use as models of PS-interacting protein function or interaction, and assays for screening candidate therapeutics preferably employ transformed eukaryotic cells.

To accomplish expression in eukaryotic cells, a wide variety of vectors have been developed and are commercially available which allow inducible (e.g., LacSwitch expression vectors, Stratagene, La Jolla, CA) or cognate (e.g., pcDNA3 vectors, Invitrogen, Chatsworth, CA) expression of PS-interacting protein nucleotide sequences under the regulation of an artificial promoter element. Such promoter elements are often derived from CMV or SV40 viral genes, although other strong promoter elements which are active in eukaryotic cells can also be employed to induce transcription of PS-interacting protein nucleotide sequences. Typically, these vectors also contain an artificial polyadenylation sequence and 3' UTR which can also be derived from exogenous viral gene sequences or from other eukaryotic genes. Furthermore, in some constructs, artificial, non-coding, spliceable introns and exons are included in the vector to enhance expression of the nucleotide sequence of interest. These expression systems are commonly available from commercial sources and are typified by vectors such as pcDNA3 and pZeoSV (Invitrogen, San Diego, CA). Innumerable commercially-available as well as custom-designed expression vectors are available from commercial sources to allow expression of any desired PS-interacting protein transcript in more or less any desired cell type, either constitutively or after exposure to a certain exogenous stimulus (e.g., withdrawal of tetracycline or exposure to IPTG).

Vectors may be introduced into the recipient or "host" cells by various methods well known in the art including, but not limited to, calcium phosphate transfection, strontium phosphate transfection, DEAE dextran transfection, electroporation, lipofection (e.g., Dosper Liposomal transfection reagent, Boehringer

Mannheim, Germany), microinjection, ballistic insertion on micro-beads, protoplast fusion or, for viral or phage vectors, by infection with the recombinant virus or phage.

5. Transgenic Animal Models

The present invention also provides for the production of transgenic non-human animal models in which mutant or wild type PS-interacting protein sequences are expressed, or in which the PS-interacting protein genes have been inactivated (e.g., "knock-out" deletions), for the study of Alzheimer's Disease, for the screening of candidate pharmaceutical compounds, for the creation of explanted mammalian CNS cell cultures (e.g., neuronal, glial, organotypic or mixed cell cultures), and for the evaluation of potential therapeutic interventions. Prior to the present invention, a partial animal model for Alzheimer's Disease existed via the insertion and over-expression of a mutant form of the human amyloid precursor protein gene as a minigene under the regulation of the platelet-derived growth factor β receptor promoter element (Games et al., 1995). This mutant (βAPP_{717} Val→Ile) causes the appearance of synaptic pathology and amyloid β peptide deposition in the brain of transgenic animals bearing this transgene in high copy number. These changes in the brain of the transgenic animal are very similar to that seen in human AD (Games et al., 1995). It is, however, as yet unclear whether these animals become demented, but there is general consensus that it is now possible to recreate at least some aspects of AD in mice. In addition, transgenic animal models in which the presenilin genes are genetically engineered are disclosed in PCT Publication WO96/34099. These transgenic animal models have been shown to have altered Aβ production and altered hippocampus-dependent memory function.

Animal species suitable for use in the animal models of the present invention include, but are not limited to, rats, mice, hamsters, guinea pigs, rabbits, dogs, cats, goats, sheep, pigs, and non-human primates (e.g., Rhesus monkeys, chimpanzees). For initial studies, transgenic rodents (e.g., mice) may be preferred due to their relative ease of maintenance and shorter life spans. However, transgenic yeast or invertebrates (e.g., nematodes, insects) may be preferred for some studies because

they will allow for even more rapid and inexpensive screening. For example, invertebrates bearing mutant PS-interacting protein homologues (or mammalian PS-interacting protein transgenes) which cause a rapidly occurring and easily scored phenotype (e.g., abnormal vulva or eye development after several days) can be used as screens for drugs which block the effect of the mutant gene. Such invertebrates may prove far more rapid and efficient for mass screenings than larger vertebrate animals. Once lead compounds are found through such screens, they may be tested in higher animals such a rodents. Ultimately, transgenic non-human primates may be preferred for longer term studies due to their greater similarity to humans and their higher cognitive abilities.

Using the nucleic acids disclosed and otherwise enabled herein, there are now several available approaches for the creation of a transgenic animal model for Alzheimer's Disease. Thus, the enabled animal models include: (1) Animals in which sequences encoding at least a functional domain of a normal human PS-interacting protein gene have been recombinantly introduced into the genome of the animal as an additional gene, under the regulation of either an exogenous or an endogenous promoter element, and as either a minigene or a large genomic fragment; in which sequences encoding at least a functional domain of a normal human PS-interacting protein gene have been recombinantly substituted for one or both copies of the animal's homologous PS-interacting protein gene by homologous recombination or gene targeting; and/or in which one or both copies of one of the animal's homologous PS-interacting protein genes have been recombinantly "humanized" by the partial substitution of sequences encoding the human homologue by homologous recombination or gene targeting. These animals are useful for evaluating the effects of the transgenic procedures, and the effects of the introduction or substitution of a human or humanized PS-interacting protein gene. (2) Animals in which sequences encoding at least a functional domain of a mutant (i.e., pathogenic) human PS-interacting protein gene have been recombinantly introduced into the genome of the animal as an additional gene, under the regulation of either an exogenous or an endogenous promoter element, and as either a minigene or a large genomic fragment;

in which sequences encoding at least a functional domain of a mutant human PS-interacting protein gene have been recombinantly substituted for one or both copies of the animal's homologous PS-interacting protein gene by homologous recombination or gene targeting; and/or in which one or both copies of one of the animal's homologous PS-interacting protein genes have been recombinantly "humanized" by the partial substitution of sequences encoding a mutant human homologue by homologous recombination or gene targeting. These animals are useful as models which will display some or all of the characteristics, whether at the biochemical, physiological and/or behavioral level, of humans carrying one or more alleles which are pathogenic of Alzheimer's Disease or other diseases associated with mutations in the PS-interacting protein genes. (3) Animals in which sequences encoding at least a functional domain of a mutant version of one of that animal's PS-interacting protein genes (bearing, for example, a specific mutation corresponding to, or similar to, one of the pathogenic mutations of the human PS-interacting proteins) have been recombinantly introduced into the genome of the animal as an additional gene, under the regulation of either an exogenous or an endogenous promoter element, and as either a minigene or a large genomic fragment; and/or in which sequences encoding at least a functional domain of a mutant version of one of that animal's PS-interacting protein genes (bearing, for example, a specific mutation corresponding to, or similar to, one of the pathogenic mutations of the human PS-interacting proteins) have been recombinantly substituted for one or both copies of the animal's homologous PS-interacting protein gene by homologous recombination or gene targeting. These animals are also useful as models which will display some or all of the characteristics, whether at the biochemical, physiological and/or behavioral level, of humans carrying one or more alleles which are pathogenic of Alzheimer's Disease. (4) "Knock-out" animals in which one or both copies of one of the animal's PS-interacting protein genes have been partially or completely deleted by homologous recombination or gene targeting, or have been inactivated by the insertion or substitution by homologous recombination or gene targeting of exogenous sequences (e.g., stop codons, lox p sites). Such animals are useful models to study the effects which loss of

PS-interacting protein gene expression may have, to evaluate whether loss of function is preferable to continued expression of mutant forms, and to examine whether other genes can be recruited to replace a mutant PS-interacting protein or to intervene with the effects of other genes (e.g., PS1, PS2, APP or ApoE) causing AD as a treatment for AD or other disorders. For example, a normal PS-interacting protein gene may be necessary for the action of mutant presenilin or APP genes to actually be expressed as AD and, therefore, transgenic PS-interacting protein animal models may be of use in elucidating such multigenic interactions.

In addition to transgenic animal models in which the expression of one or more of the PS-interacting proteins is altered, the present invention also provides for the production of transgenic animal models in which the expression of one or more of the presenilins, APP, or ApoE is altered. The nucleic acids encoding the presenilins, APP, and ApoE are known in the art, a methods for producing transgenic animals with these sequences are also known (see, e.g., PCT Publication WO96/34099; Games et al., 1995). Indeed, because non-human animals may differ from humans not only in their PS-interacting protein sequences, but also in the sequences of their presenilin, APP and/or ApoE homologues, it is particularly contemplated that transgenics may be produced which bear recombinant normal or mutant human sequences for at least one presenilin, APP and/or ApoE gene in addition to recombinant sequences for one or more PS-interacting proteins. Such co-transformed animal models would possess more elements of the human molecular biology and, therefore, are expected to be better models of human disorders. Thus, in accordance with the present invention, transgenic animal models may be produced bearing normal or mutant sequences for one or more PS-interacting proteins, or interacting domains of these proteins. These animals will have utility in that they can be crossed with animals bearing a variety of normal or mutant presenilin, APP or ApoE sequences to produce co-transformed animal models. Furthermore, as detailed below, it is expected that mutations in the PS-interacting genes, like mutations in the presenilins themselves, may be causative of Alzheimer's Disease and/or other disorders as well (e.g., other cognitive, intellectual, neurological or psychological disorders such as cerebral hemorrhage,

schizophrenia, depression, mental retardation and epilepsy). Therefore, transgenic animal models bearing normal or mutant sequences corresponding to the PS-interacting proteins, absent transformation with any presenilin, APP or ApoE sequences, will have utility of their own in the study of such disorders.

As detailed below, preferred choices for transgenic animal models transformed with PS-interacting proteins, or domains of PS-interacting proteins, include those transformed with normal or mutant sequences corresponding to the clones identified and described in Example 1 and disclosed in SEQ ID NOs: 1-12. These clones, which interact with normal or mutant PS1 TM6→7 loop domains, were identified according to the methods described in Example 1, below, and PCT Publication WO96/34099. These clones, longer nucleic acid sequences comprising these clones, and other clones identified according to this and other methods of the invention (e.g., allelic and splice variants or heterospecific homologues of these clones) may all be employed in accordance with the present invention to produce animal models which, with or without co-transformation with presenilin, APP and/or ApoE sequences, will have utility in the study of Alzheimer's Disease and/or other cognitive, intellectual, neurological or psychological disorders.

Thus, using the nucleic acids disclosed and otherwise enabled herein, one of ordinary skill in the art may now produce any of the following types of transgenic animal models with altered PS-interacting protein expression: (1) Animals in which sequences encoding at least a functional domain of a normal human PS-interacting protein gene have been recombinantly introduced into the genome of the animal as an additional gene, under the regulation of either an exogenous or an endogenous promoter element, and as either a minigene or a large genomic fragment; in which sequences encoding at least a functional domain of a normal human PS-interacting protein gene have been recombinantly substituted for one or both copies of the animal's homologous PS-interacting protein gene by homologous recombination or gene targeting; and/or in which one or both copies of one of the animal's homologous PS-interacting protein genes have been recombinantly "humanized" by the partial substitution of sequences encoding the human homologue by homologous

recombination or gene targeting. These animals are particularly useful for providing transgenic models which express human PS-interacting proteins as well as human presenilin proteins. They are also useful in evaluating the effects of the transgenic procedures, and the effects of the introduction or substitution of a human or humanized PS-interacting protein gene. (2) Animals in which sequences encoding at least a functional domain of a mutant (i.e., pathogenic) human PS-interacting protein gene have been recombinantly introduced into the genome of the animal as an additional gene, under the regulation of either an exogenous or an endogenous promoter element, and as either a minigene or a large genomic fragment; in which sequences encoding at least a functional domain of a mutant human PS-interacting protein gene have been recombinantly substituted for one or both copies of the animal's homologous PS-interacting protein gene by homologous recombination or gene targeting; and/or in which one or both copies of one of the animal's homologous PS-interacting protein genes have been recombinantly "humanized" by the partial substitution of sequences encoding a mutant human homologue by homologous recombination or gene targeting. These animals are useful as models which will display some or all of the characteristics, whether at the biochemical, physiological and/or behavioral level, of humans carrying one or more alleles which are pathogenic of Alzheimer's Disease or other diseases associated with mutations in these PS-interacting genes. (3) Animals in which sequences encoding at least a functional domain of a mutant version of one of that animal's PS-interacting protein genes (bearing, for example, a specific mutation corresponding to, or similar to, one of the pathogenic mutations of the human PS-interacting proteins) have been recombinantly introduced into the genome of the animal as an additional gene, under the regulation of either an exogenous or an endogenous promoter element, and as either a minigene or a large genomic fragment; and/or in which sequences encoding at least a functional domain of a mutant version of one of that animal's PS-interacting protein genes (bearing, for example, a specific mutation corresponding to, or similar to, one of the pathogenic mutations of the humans PS-interacting proteins) have been recombinantly substituted for one or both copies of the animal's homologous PS-interacting protein

gene by homologous recombination or gene targeting. These animals are also useful as models which will display some or all of the characteristics, whether at the biochemical, physiological and/or behavioral level, of humans carrying one or more alleles which are pathogenic of Alzheimer's Disease. (4) "Knock-out" animals in which one or both copies of one of the animal's PS-interacting protein genes have been partially or completely deleted by homologous recombination or gene targeting, or have been inactivated by the insertion or substitution by homologous recombination or gene targeting of exogenous sequences (e.g., stop codons, lox p sites). Such animals are useful models to study the effects which loss of PS-interacting protein gene expression may have, to evaluate whether loss of function is preferable to continued expression, and to examine whether other genes can be recruited to replace a mutant PS-interacting protein or to intervene with the effects of other genes (e.g., APP or ApoE) causing AD as a treatment for AD or other disorders. For example, a normal PS-interacting protein may be necessary for the action of mutant PS1, PS2 or APP genes to actually be expressed as AD and, therefore, transgenic PS-interacting protein animal models may be of use in elucidating such multigenic interactions.

In some preferred embodiments, transgenic animal models are produced in which just the PS-interacting domains of the PS-interacting proteins are introduced into the genome of the animal by homologous recombination. Thus, for example, preferred embodiments include transgenic animals in which the PS-interacting domains of PS-interacting proteins are "humanized" by homologous recombination with sequences from human PS-interacting proteins. These animals may then be bred with transgenics in which normal or mutant presenilin sequences have been introduced. The progeny of these animals, having both human presenilin and human PS-interacting protein sequences, will provide improved animal models for Alzheimer's Disease.

To create an animal model (e.g., a transgenic mouse), a normal or mutant PS-interacting gene (e.g., normal or mutant S5a, GT24, p0071, Rab11, etc.), or a normal or mutant version of a recombinant nucleic acid encoding at least a functional

domain of a PS-interacting gene (e.g., the PS-interacting domains obtained in the yeast two-hybrid system), can be inserted into a germ line or stem cell using standard techniques of oocyte microinjection, or transfection or microinjection into embryonic stem cells. Animals produced by these or similar processes are referred to as transgenic. Similarly, if it is desired to inactivate or replace an endogenous presenilin or PS-interacting protein gene, homologous recombination using embryonic stem cells may be employed. Animals produced by these or similar processes are referred to as "knock-out" (inactivation) or "knock-in" (replacement) models.

For oocyte injection, one or more copies of the recombinant DNA constructs of the present invention may be inserted into the pronucleus of a just-fertilized oocyte. This oocyte is then reimplanted into a pseudo-pregnant foster mother. The liveborn animals are screened for integrants using analysis of DNA (e.g., from the tail veins of offspring mice) for the presence of the inserted recombinant transgene sequences. The transgene may be either a complete genomic sequence injected as a YAC, BAC, PAC or other chromosome DNA fragment, a cDNA with either the natural promoter or a heterologous promoter, or a minigene containing all of the coding region and other elements found to be necessary for optimum expression.

Retroviral infection of early embryos can also be done to insert the recombinant DNA constructs of the invention. In this method, the transgene (e.g., a normal or mutant S5a, GT24, p0071, Rab 11, etc., sequence) is inserted into a retroviral vector which is used to infect embryos (e.g., mouse or non-human primate embryos) directly during the early stages of development to generate chimeras, some of which will lead to germline transmission.

Homologous recombination using stem cells allows for the screening of gene transfer cells to identify the rare homologous recombination events. Once identified, these can be used to generate chimeras by injection of blastocysts, and a proportion of the resulting animals will show germline transmission from the recombinant line. This methodology is especially useful if inactivation of a gene is desired. For example, inactivation of the S5a gene in mice may be accomplished by designing a DNA fragment which contains sequences from an S5a coding region

flanking a selectable marker. Homologous recombination leads to the insertion of the marker sequences in the middle of the coding region, causing inactivation of the S5a gene and/or deletion of internal sequences. DNA analysis of individual clones can then be used to recognize the homologous recombination events.

The techniques of generating transgenic animals, as well as the techniques for homologous recombination or gene targeting, are now widely accepted and practiced. A laboratory manual on the manipulation of the mouse embryo, for example, is available detailing standard laboratory techniques for the production of transgenic mice (Hogan et al., 1986). To create a transgene, the target sequence of interest (e.g., normal or mutant presenilin sequences, normal or mutant PS-interacting protein sequences) are typically ligated into a cloning site located downstream of some promoter element which will regulate the expression of RNA from the sequence. Downstream of the coding sequence, there is typically an artificial polyadenylation sequence. In the transgenic models that have been used to successfully create animals which mimic aspects of inherited human neurodegenerative diseases, the most successful promoter elements have been the platelet-derived growth factor receptor β gene subunit promoter and the hamster prion protein gene promoter, although other promoter elements which direct expression in central nervous system cells would also be useful. An alternate approach to creating a transgene is to use an endogenous presenilin or PS-interacting protein gene promoter and regulatory sequences to drive expression of the transgene. Finally, it is possible to create transgenes using large genomic DNA fragments such as YACs which contain the entire desired gene as well as its appropriate regulatory sequences. Such constructs have been successfully used to drive human APP expression in transgenic mice (Lamb et al., 1993).

Animal models can also be created by targeting the endogenous presenilin or PS-interacting protein gene in order to alter the endogenous sequence by homologous recombination. These targeting events can have the effect of removing endogenous sequence (knock-out) or altering the endogenous sequence to create an amino acid change associated with human disease or an otherwise abnormal sequence

(e.g., a sequence which is more like the human sequence than the original animal sequence) (knock-in animal models). A large number of vectors are available to accomplish this and appropriate sources of genomic DNA for mouse and other animal genomes to be targeted are commercially available from companies such as GenomeSystems Inc. (St. Louis, Missouri, USA). The typical feature of these targeting vector constructs is that 2 to 4 kb of genomic DNA is ligated 5' to a selectable marker (e.g., a bacterial neomycin resistance gene under its own promoter element termed a "neomycin cassette"). A second DNA fragment from the gene of interest is then ligated downstream of the neomycin cassette but upstream of a second selectable marker (e.g., thymidine kinase). The DNA fragments are chosen such that mutant sequences can be introduced into the germ line of the targeted animal by homologous replacement of the endogenous sequences by either one of the sequences included in the vector. Alternatively, the sequences can be chosen to cause deletion of sequences that would normally reside between the left and right arms of the vector surrounding the neomycin cassette. The former is known as a knock-in, the latter is known as a knock-out. Again, innumerable model systems have been created, particularly for targeted knock-outs of genes including those relevant to neurodegenerative diseases (e.g., targeted deletions of the murine APP gene by Zheng et al., 1995; targeted deletion of the murine prion gene associated with adult onset human CNS degeneration by Bueler et al., 1996).

Finally, equivalents of transgenic animals, including animals with mutated or inactivated presenilin genes, or mutated or inactivated PS-interacting protein genes, may be produced using chemical or X-ray mutagenesis of gametes, followed by fertilization. Using the isolated nucleic acids disclosed or otherwise enabled herein, one of ordinary skill may more rapidly screen the resulting offspring by, for example, direct sequencing RFLP, PCR, or hybridization analysis to detect mutants, or Southern blotting to demonstrate loss of one allele by dosage.

6. Assays for Drugs Which Affect PS-Interacting Protein Expression

In another series of embodiments, the present invention provides assays for identifying small molecules or other compounds which are capable of inducing or inhibiting the expression of the PS-interacting genes and proteins (e.g., S5a or GT24). The assays may be performed in vitro using non-transformed cells, immortalized cell lines, or recombinant cell lines, or in vivo using the transgenic animal models enabled herein.

In particular, the assays may detect the presence of increased or decreased expression of S5a, GT24, p0071, Rab 11, or other PS-interacting genes or proteins on the basis of increased or decreased mRNA expression (using, e.g., the nucleic acid probes disclosed and enabled herein), increased or decreased levels of PS-interacting proteins (using, e.g., the anti-PS-interacting protein antibodies disclosed and enabled herein), or increased or decreased levels of expression of a marker gene (e.g., β-galactosidase or luciferase) operably joined to a PS-interacting protein 5' regulatory region in a recombinant construct.

Thus, for example, one may culture cells known to express a particular PS-interacting protein and add to the culture medium one or more test compounds. After allowing a sufficient period of time (e.g., 0-72 hours) for the compound to induce or inhibit the expression of the PS-interacting protein, any change in levels of expression from an established baseline may be detected using any of the techniques described above and well known in the art. In particularly preferred embodiments, the cells are from an immortalized cell line such as a human neuroblastoma, glioblastoma or a hybridoma cell line. Using the nucleic acid probes and /or antibodies disclosed and enabled herein, detection of changes in the expression of a PS-interacting protein, and thus identification of the compound as an inducer or repressor of PS-interacting protein expression, requires only routine experimentation.

In particularly preferred embodiments, a recombinant assay is employed in which a reporter gene such a β-galactosidase, green fluorescent protein , alkaline phosphatase, or luciferase is operably joined to the 5' regulatory regions of a PS-interacting protein gene. Preferred vectors include the Green Lantern 1 vector (GIBCO/BRL, Gaithersburg, MD) and the Great EScAPe pSEAP vector (Clontech,

Palo Alto). The PS-interacting protein regulatory regions may be easily isolated and cloned by one of ordinary skill in the art in light of the present disclosure of coding regions from these genes. The reporter gene and regulatory regions are joined in-frame (or in each of the three possible reading frames) so that transcription and translation of the reporter gene may proceed under the control of the PS-interacting protein regulatory elements. The recombinant construct may then be introduced into any appropriate cell type, although mammalian cells are preferred, and human cells are most preferred. The transformed cells may be grown in culture and, after establishing the baseline level of expression of the reporter gene, test compounds may be added to the medium. The ease of detection of the expression of the reporter gene provides for a rapid, high through-put assay for the identification of inducers and repressors of the PS-interacting protein gene.

Compounds identified by this method will have potential utility in modifying the expression of the PS-interacting protein genes in vivo. These compounds may be further tested in the animal models disclosed and enabled herein to identify those compounds having the most potent in vivo effects. In addition, as described herein with respect to small molecules having binding activity for PS-interacting proteins, these molecules may serve as "lead compounds" for the further development of pharmaceuticals by, for example, subjecting the compounds to sequential modifications, molecular modeling, and other routine procedures employed in rational drug design.

7. Identification of Compounds with PS-Interacting Protein Binding Capacity

In light of the present disclosure, one of ordinary skill in the art is enabled to practice new screening methodologies which will be useful in the identification of proteins and other compounds which bind to, or otherwise directly interact with, the PS-interacting proteins. The proteins and compounds will include endogenous cellular components, aside from the presenilins, which interact with the PS-interacting proteins in vivo and which, therefore, provide new targets for pharmaceutical and therapeutic interventions, as well as recombinant, synthetic and otherwise exogenous

compounds which may have PS-interacting protein binding capacity and, therefore, may be candidates for pharmaceutical agents. Thus, in one series of embodiments, cell lysates or tissue homogenates (e.g., human brain homogenates, lymphocyte lysates) may be screened for proteins or other compounds which bind to one of the normal or mutant PS-interacting proteins. Alternatively, any of a variety of exogenous compounds, both naturally occurring and/or synthetic (e.g., libraries of small molecules or peptides), may be screened for PS-interacting protein binding capacity. Small molecules are particular preferred in this context because they are more readily absorbed after oral administration, have fewer potential antigenic determinants, and/or are more likely to cross the blood brain barrier than larger molecules such as nucleic acids or proteins. The methods of the present invention are particularly useful in that they may be used to identify molecules which selectively or preferentially bind to a mutant form of a PS-interacting protein (rather than a normal form) and, therefore, may have particular utility in treating cases of AD which arise from mutations in the PS-interacting proteins.

Once identified by the methods described above, the candidate compounds may then be produced in quantities sufficient for pharmaceutical administration or testing (e.g., μg or mg or greater quantities), and formulated in a pharmaceutically acceptable carrier (see, e.g., Remington's Pharmaceutical Sciences, Gennaro, A., ed., Mack Pub., 1990). These candidate compounds may then be administered to the transformed cells of the invention, to the transgenic animal models of the invention, to cell lines derived from the animal models or from human patients, or to Alzheimer's patients. The animal models described and enabled herein are of particular utility in further testing candidate compounds which bind to normal or mutant PS-interacting proteins for their therapeutic efficacy.

In addition, once identified by the methods described above, the candidate compounds may also serve as "lead compounds" in the design and development of new pharmaceuticals. For example, as in well known in the art, sequential modification of small molecules (e.g., amino acid residue replacement with peptides; functional group replacement with peptide or non-peptide compounds) is a standard

approach in the pharmaceutical industry for the development of new pharmaceuticals. Such development generally proceeds from a "lead compound" which is shown to have at least some of the activity (e.g., PS-interacting protein binding or blocking ability) of the desired pharmaceutical. In particular, when one or more compounds having at least some activity of interest (e.g., modulation of PS-interacting protein activity) are identified, structural comparison of the molecules can greatly inform the skilled practitioner by suggesting portions of the lead compounds which should be conserved and portions which may be varied in the design of new candidate compounds. Thus, the present invention also provides a means of identifying lead compounds which may be sequentially modified to produce new candidate compounds for use in the treatment of Alzheimer's Disease. These new compounds then may be tested both for binding to PS-interacting proteins and/or blocking PS-interacting protein activity, and for therapeutic efficacy (e.g., in the animal models described herein). This procedure may be iterated until compounds having the desired therapeutic activity and/or efficacy are identified.

In each of the present series of embodiments, an assay is conducted to detect binding between a "PS-interacting protein component" and some other moiety. Of particular utility will be sequential assays in which compounds are tested for the ability to bind to only normal or only mutant forms of the PS-interacting domains of PS-interacting proteins in the binding assays. Such compounds are expected to have the greatest therapeutic utilities, as described more fully below. The "PS-interacting protein component" in these assays may be a complete normal or mutant form of a PS-interacting protein (e.g., S5a, GT24, p0071, Rab 11, etc.) but need not be. Rather, particular functional domains of the PS-interacting proteins, particularly the PS-interacting domains as described above, may be employed either as separate molecules or as part of a fusion protein. For example, to isolate proteins or compounds that interact with these functional domains, screening may be carried out using fusion constructs and/or synthetic peptides corresponding to these regions. Thus, for S5a, GST-fusion peptides may be made including sequences corresponding approximately to amino acids 70-377 of SEQ ID NO: 2 (included in clones Y2H29

and Y2H31, see Example 1), approximately to amino acids 206-377 of SEQ ID NO: 2 (which includes protein-protein interaction motifs, see Ferrell et al., 1996), or to any other S5a domain of interest. Similarly, for GT24, GST- or other fusion peptides may be produced including sequences corresponding approximately to amino acids 440-815 of SEQ ID NO: 4 (including part of the armadillo repeat segment). Obviously, various combinations of fusion proteins and PS-interacting protein functional domains are possible and these are merely examples. In addition, the functional domains may be altered so as to aid in the assay by, for example, introducing into the functional domain a reactive group or amino acid residue (e.g., cysteine) which will facilitate immobilization of the domain on a substrate (e.g., using sulfhydryl reactions). Thus, for example, the PS-interacting domain of S5a may be synthesized containing an additional C-terminal cysteine residue to facilitate immobilization of the domain. Such peptides may be used to create an affinity substrate for affinity chromatography (Sulfo-link; Pierce) to isolate binding proteins for microsequencing. Similarly, other functional domain or antigenic fragments may be created with modified residues (see, e.g., Example 4).

The proteins or other compounds identified by these methods may be purified and characterized by any of the standard methods known in the art. Proteins may, for example, be purified and separated using electrophoretic (e.g., SDS-PAGE, 2D PAGE) or chromatographic (e.g., HPLC) techniques and may then be microsequenced. For proteins with a blocked N-terminus, cleavage (e.g., by CNBr and/or trypsin) of the particular binding protein is used to release peptide fragments. Further purification/characterization by HPLC and microsequencing and/or mass spectrometry by conventional methods provides internal sequence data on such blocked proteins. For non-protein compounds, standard organic chemical analysis techniques (e.g., IR, NMR and mass spectrometry; functional group analysis; X-ray crystallography) may be employed to determine their structure and identity.

Methods for screening cellular lysates, tissue homogenates, or small molecule libraries for candidate PS-interaction protein-binding molecules are well known in the art and, in light of the present disclosure, may now be employed to

identify compounds which bind to normal or mutant PS-interacting protein components or which modulate PS-interacting protein activity as defined by non-specific measures (e.g., changes in intracellular Ca^{2+}, GTP/GDP ratio) or by specific measures (e.g., changes in Aβ peptide production or changes in the expression of other downstream genes which can be monitored by differential display, 2D gel electrophoresis, differential hybridization, or SAGE methods). The preferred methods involve variations on the following techniques: (1) direct extraction by affinity chromatography; (2) co-isolation of PS-interacting protein components and bound proteins or other compounds by immunoprecipitation; (3) the Biomolecular Interaction Assay (BIAcore); and (4) the yeast two-hybrid systems. These and others are discussed separately below.

A. Affinity Chromatography

In light of the present disclosure, a variety of affinity binding techniques well known in the art may be employed to isolate proteins or other compounds which bind to the PS-interacting protein disclosed or otherwise enabled herein. In general, a PS-interacting protein component may be immobilized on a substrate (e.g., a column or filter) and a solution including the test compound(s) is contacted with the PS-interacting protein, fusion or fragment under conditions which are permissive for binding. The substrate is then washed with a solution to remove unbound or weakly bound molecules. A second wash may then elute those compounds which strongly bound to the immobilized normal or mutant PS-interacting protein component. Alternatively, the test compounds may be immobilized and a solution containing one or more PS-interacting protein components may be contacted with the column, filter or other substrate. The ability of the PS-interacting protein component to bind to the test compounds may be determined as above or a labeled form of the PS-interacting protein component (e.g., a radio-labeled or chemiluminescent functional domain) may be used to more rapidly assess binding to the substrate-immobilized compound(s).

B. Co-Immunoprecipitation

Another well characterized technique for the isolation of PS-interacting protein components and their associated proteins or other compounds is direct

immunoprecipitation with antibodies. This procedure has been successfully used, for example, to isolate many of the synaptic vesicle associated proteins (Phizicky and Fields, 1994). Thus, either normal or mutant, free or membrane-bound PS-interacting protein components may be mixed in a solution with the candidate compound(s) under conditions which are permissive for binding, and the PS-interacting protein component may be immunoprecipitated. Proteins or other compounds which co-immunoprecipitate with the PS-interacting protein component may then be identified by standard techniques as described above. General techniques for immunoprecipitation may be found in, for example, Harlow and Lane, (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY.

The antibodies employed in this assay, as described and enabled herein, may be polyclonal or monoclonal, and include the various antibody fragments (e.g., Fab, $F(ab')_2$,) as well as single chain antibodies, and the like.

C. The Biomolecular Interaction Assay

Another useful method for the detection and isolation of binding proteins is the Biomolecular Interaction Assay or "BIAcore" system developed by Pharmacia Biosensor and described in the manufacturer's protocol (LKB Pharmacia, Sweden). In light of the present disclosure, one of ordinary skill in the art is now enabled to employ this system, or a substantial equivalent, to identify proteins or other compounds having PS-interacting protein binding capacity. The BIAcore system uses an affinity purified anti-GST antibody to immobilize GST-fusion proteins onto a sensor chip. Obviously, other fusion proteins and corresponding antibodies may be substituted. The sensor utilizes surface plasmon resonance which is an optical phenomenon that detects changes in refractive indices. A homogenate of a tissue of interest is passed over the immobilized fusion protein and protein-protein interactions are registered as changes in the refractive index. This system can be used to determine the kinetics of binding and to assess whether any observed binding is of physiological relevance.

D. The Yeast Two-Hybrid System

The yeast "two-hybrid" system takes advantage of transcriptional factors that are composed of two physically separable, functional domains (Phizicky and Fields, 1994). The most commonly used is the yeast GAL4 transcriptional activator consisting of a DNA binding domain and a transcriptional activation domain. Two different cloning vectors are used to generate separate fusions of the GAL4 domains to genes encoding potential binding proteins. The fusion proteins are co-expressed, targeted to the nucleus and, if interactions occur, activation of a reporter gene (e.g., lacZ) produces a detectable phenotype. For example, the Clontech Matchmaker System-2 may be used with the Clontech brain cDNA GAL4 activation domain fusion library with PS-interacting protein-GAL4 binding domain fusion clones (Clontech, Palo Alto, CA). In light of the disclosures herein, one of ordinary skill in the art is now enabled to produce a variety of PS-interacting protein fusions, including fusions including either normal or mutant functional domains of the PS-interacting proteins, and to screen such fusion libraries in order to identify PS-interacting protein binding proteins.

E. Other Methods

The nucleotide sequences and protein products, including both mutant and normal forms of these nucleic acids and their corresponding proteins, can be used with the above techniques to isolate other interacting proteins, and to identify other genes whose expression is altered by the over-expression of normal PS-interacting protein sequences, by the under-expression of normal PS-interacting protein sequences, or by the expression of mutant PS-interacting protein sequences. Identification of these other interacting proteins, as well as the identification of other genes whose expression levels are altered in AD will identify other gene targets which have direct relevance to the pathogenesis of this disease in its clinical or pathological forms. Specifically, other genes will be identified which may themselves be the site of other mutations causing Alzheimer's Disease, or which can themselves be targeted therapeutically (e.g., to reduce their expression levels to normal, or to pharmacologically block the effects of their over-expression) as a potential treatment

for this disease. Specifically, these techniques rely on PCR-based and/or hybridization-based methods to identify genes which are differentially expressed between two conditions (a cell line expressing normal PS-interacting proteins compared to the same cell type expressing a mutant PS-interacting protein). These techniques include differential display, serial analysis of gene expression (SAGE), and mass-spectrometry of protein 2D-gels and subtractive hybridization (reviewed in Nowak, 1995 and Kahn, 1995).

As will be obvious to one of ordinary skill in the art, there are numerous other methods of screening individual proteins or other compounds, as well as large libraries of proteins or other compounds (e.g., phage display libraries and cloning systems from Stratagene, La Jolla, CA) to identify molecules which bind to normal or mutant PS-interacting protein components. All of these methods comprise the step of mixing a normal or mutant PS-interacting protein, fusion, or fragment with test compounds, allowing for binding (if any), and assaying for bound complexes. All such methods are now enabled by the present disclosure of substantially pure PS-interacting proteins, substantially pure PS-interacting functional domain fragments, PS-interacting protein fusion proteins, PS-interacting protein antibodies, and methods of making and using the same.

8. Disrupting PS-Interacting Protein Interactions

The ability to disrupt specific interactions of the PS-interacting proteins with the presenilins, or with other proteins, is potentially of great therapeutic value, and will be important in understanding the etiology of AD and in identifying additional targets for therapy. The methods used to identify compounds which disrupt PS-interacting protein interactions may be applied equally well to interactions involving either normal or mutant PS-interacting proteins.

Assays for compounds which can disrupt PS-interacting protein interactions may be performed by any of a variety of methods well known in the art. In essence, such assays will parallel those assays for identifying proteins and compounds with binding activity toward the PS-interacting proteins. Thus, once a

compound with binding activity for a PS-interacting protein is identified by any method, that method or an equivalent method may be performed in the presence of candidate compounds to identify compounds which disrupt the interaction. Thus, for example, the assay may employ methods including (1) affinity chromatography; (2) immunoprecipitation; (3) the Biomolecular Interaction Assay (BIAcore); or (4) the yeast two-hybrid systems. Such assays can be developed using either normal or mutant purified PS-interacting proteins, and/or either normal or mutant purified binding proteins (e.g., normal or mutant presenilins).

For affinity methods, either the PS-interacting protein or its binding partner may be affixed to a matrix, for example in a column, and the counterpart protein (e.g., the PS-interacting protein if presenilin or another binding partner is affixed to the matrix; or a presenilin or other binding partner if the PS-interacting protein is affixed to the matrix) is then exposed to the affixed protein/compound either before or after adding the candidate compound(s). In the absence of a disruptive effect by the candidate compound(s), the interaction between the PS-interacting protein and its binding partner will cause the counterpart protein to bind to the affixed protein. Any compound which disrupts the interaction will cause release of the counterpart protein from the matrix. Release of the counterpart protein from the matrix can be measured using methods known in the art.

For PS-interacting protein interactions which are detectable by yeast two-hybrid systems, these assays may also be employed to identify compounds which disrupt the interaction. Briefly, a PS-interacting protein and its binding partner (or appropriate structural domains of each) are employed in the fusion proteins of the system, and the cells are exposed to candidate compounds to determine their effect upon the expression of the reporter gene. By appropriate choice of a reporter gene, such a system can be readily adapted for high through-put screening of large libraries of compounds by, for example, using a reporter gene which confers resistance to an antibiotic which is present in the medium, or which rescues an auxotrophic strain grown in minimal medium.

These assays may be used to screen many different types of compounds for their disruptive effect on the interactions of the PS-interacting proteins. For example, the compounds may belong to a library of synthetic molecules, or be specifically designed to disrupt the interaction. The compounds may also be peptides corresponding to the interacting domain of either protein. This type of assay can be used to identify compounds that disrupt a specific interaction between a given PS-interacting protein variant and a given binding partner. In addition, compounds that disrupt all interactions with PS-interacting proteins may be identified. For example, a compound that specifically disrupts the folding of PS-interacting proteins would be expected to disrupt all interactions between PS-interacting proteins and other proteins. Alternatively, this type of disruption assay can be used to identify compounds which disrupt only a range of different PS-interacting protein interactions, or only a single PS-interacting protein interaction.

9. Methods of Identifying Compounds Modulating PS-Interacting Protein Activity

In another series of embodiments, the present invention provides for methods of identifying compounds with the ability to modulate the activity of normal and mutant PS-interacting proteins. As used with respect to this series of embodiments, the term "activity" broadly includes gene and protein expression, PS-interacting protein post-translation processing, trafficking and localization, and any functional activity (e.g., enzymatic, receptor-effector, binding, channel), as well as downstream affects of any of these. It is known that Alzheimer's Disease is associated with increased production of the long form of Aβ peptides, the appearance of amyloid plaques and neurofibrillary tangles, decreases in cognitive function, and apoptotic cell death. Therefore, using the transformed cells and transgenic animal models of the present invention, cells obtained from subjects bearing normal or mutant PS-interacting protein genes, or animals or human subjects bearing naturally occurring normal or mutant PS-interacting proteins, it is now possible to screen candidate pharmaceuticals and treatments for their therapeutic effects by detecting changes in

one or more of these functional characteristics or phenotypic manifestations of normal or mutant PS-interacting protein expression.

Thus, the present invention provides methods for screening or assaying for proteins, small molecules or other compounds which modulate PS-interacting protein activity by contacting a cell in vivo or in vitro with a candidate compound and assaying for a change in a marker associated with normal or mutant PS-interacting protein activity. The marker associated with PS-interacting protein activity may be any measurable biochemical, physiological, histological and/or behavioral characteristic associated with PS-interacting protein expression. In particular, useful markers will include any measurable biochemical, physiological, histological and/or behavioral characteristic which distinguishes cells, tissues, animals or individuals bearing at least one mutant presenilin or PS-interacting protein gene from their normal counterparts. In addition, the marker may be any specific or non-specific measure of presenilin or PS-interacting protein activity. PS-interacting protein specific measures include measures of PS-interacting protein expression (e.g., PS-interacting protein mRNA or protein levels) which may employ the nucleic acid probes or antibodies of the present invention. Non-specific measures include changes in cell physiology such as pH, intracellular calcium, cyclic AMP levels, GTP/GDP ratios, phosphatidylinositol activity, protein phosphorylation, etc., which can be monitored on devices such as the cytosensor microphysiometer (Molecular Devices Inc., United States). The activation or inhibition of PS-interacting protein activity in its mutant or normal form can also be monitored by examining changes in the expression of other genes (e.g., the presenilins) which are specific to the PS-interacting protein pathway leading to Alzheimer's Disease. These can be assayed by such techniques as differential display, differential hybridization, and SAGE (sequential analysis of gene expression), as well as by two dimensional gel electrophoresis of cellular lysates. In each case, the differentially-expressed genes can be ascertained by inspection of identical studies before and after application of the candidate compound. Furthermore, as noted elsewhere, the particular genes whose expression is modulated by the administration of the candidate compound can be ascertained by cloning,

nucleotide sequencing, amino acid sequencing, or mass spectrometry (reviewed in Nowak, 1995).

In general, a cell may be contacted with a candidate compound and, after an appropriate period (e.g., 0-72 hours for most biochemical measures of cultured cells), the marker of presenilin or PS-interacting protein activity may be assayed and compared to a baseline measurement. The baseline measurement may be made prior to contacting the cell with the candidate compound or may be an external baseline established by other experiments or known in the art. The cell may be a transformed cell of the present invention or an explant from an animal or individual. In particular, the cell may be an explant from a carrier of a presenilin or PS-interacting protein mutation (e.g., a human subject with Alzheimer's Disease) or an animal model of the invention (e.g., a transgenic nematode or mouse bearing a mutant presenilin or PS-interacting protein gene). To augment the effect of presenilin or PS-interacting protein mutations on the Aβ pathway, transgenic cells or animals may be employed which have increased Aβ production. Preferred cells include those from neurological tissues such as neuronal, glial or mixed cell cultures; and cultured fibroblasts, liver, kidney, spleen, or bone marrow. The cells may be contacted with the candidate compounds in a culture in vitro or may be administered in vivo to a live animal or human subject. For live animals or human subjects, the test compound may be administered orally or by any parenteral route suitable to the compound. For clinical trials of human subjects, measurements may be conducted periodically (e.g., daily, weekly or monthly) for several months or years.

Because most individuals bearing a mutation in a particular gene are heterozygous at that locus (i.e., bearing one normal and one mutant allele), compounds may be tested for their ability to modulate normal as well as mutant presenilin or PS-interacting protein activity. Thus, for example, compounds which enhance the function of normal presenilins or PS-interacting proteins may have utility in treating Alzheimer's Disease or related disorders. Alternatively, because suppression of the activity of both normal and mutant copies of a gene in a heterozygous individual may have less severe clinical consequences than progression

of the associated disease, it may be desired to identify compound which inactivate or suppress all forms of the presenilins, the PS-interacting proteins, or their interactions. Preferably, however, compounds are identified which selectively or specifically inactivate or suppress the activity of mutant presenilin or PS-interacting proteins without disrupting the function of their normal counterparts.

In light of the identification, characterization, and disclosure herein of a novel group of PS-interacting genes and proteins, the PS-interacting protein nucleic acid probes and antibodies, and the PS-interacting protein transformed cells and transgenic animals of the invention, one of ordinary skill in the art is now enabled by perform a great variety of assays which will detect the modulation of presenilin and/or PS-interacting protein activity by candidate compounds. Particularly preferred and contemplated embodiments are discussed in some detail below.

A. PS-Interacting Protein Expression

In one series of embodiments, specific measures of PS-interacting protein expression are employed to screen candidate compounds for their ability to affect presenilin activity. Thus, using the PS-interacting protein nucleic acids and antibodies disclosed and otherwise enabled herein, one may use mRNA levels or protein levels as a marker for the ability of a candidate compound to modulate PS-interacting protein activity. The use of such probes and antibodies to measure gene and protein expression is well known in the art and discussed elsewhere herein. Of particular interest may be the identification of compounds which can alter the relative levels of different variants (e.g., mutant and normal) of the PS-interacting proteins.

B. Intracellular Localization

In another series of embodiments, compounds may be screened for their ability to modulate the activity of the PS-interacting proteins based upon their effects on the trafficking and intracellular localization of the PS-interacting proteins. The presenilins and some of the PS-interacting proteins (e.g., S5a) have been seen immunocytochemically to be localized in membrane structures associated with the endoplasmic reticulum and Golgi apparatus. Differences in localization of mutant and normal presenilins or PS-interacting proteins may, therefore, contribute to the etiology

of Alzheimer's Disease and related disorders. Compounds which can affect the localization of these proteins may, therefore, be identified as potential therapeutics. Standard techniques known in the art may be employed to detect the localization of the presenilins and PS-interacting proteins. Generally, these techniques will employ the antibodies of the present invention, and in particular antibodies which selectively bind to one or more mutant PS-interacting proteins but not to normal proteins. As is well known in the art, such antibodies may be labeled by any of a variety of techniques (e.g., fluorescent or radioactive tags, labeled secondary antibodies, avidin-biotin, etc.) to aid in visualizing the intracellular location of these proteins. The PS-interacting proteins may be co-localized to particular structures, as in known in the art, using antibodies to markers of those structures (e.g., TGN38 for the Golgi, transferrin receptor for post-Golgi transport vesicles, LAMP2 for lysosomes). Western blots of purified fractions from cell lysates enriched for different intracellular membrane bound organelles (e.g., lysosomes, synaptosomes, Golgi) may also be employed.

B. Ion Regulation/Metabolism

In another series of embodiments, compounds may be screened for their ability to modulate the activity of the presenilins or PS-interacting proteins based upon measures in intracellular Ca^{2+}, Na^{+} or K^{+} levels or metabolism. As noted above, the presenilins are membrane associated proteins which may serve as, or interact with, ion receptors or ion channels. Thus, compounds may be screened for their ability to modulate presenilin and PS-interacting protein-related metabolism of calcium or other ions either in vivo or in vitro by, for example, measurements of ion channel fluxes and/or transmembrane voltage and/or current fluxes, using patch clamps, voltage clamps or fluorescent dyes sensitive to intracellular ion levels or transmembrane voltage. Ion channel or receptor function can also be assayed by measurements of activation of second messengers such as cyclic AMP, cGMP tyrosine kinases, phosphates, increases in intracellular Ca^{2+} levels, etc. Recombinantly made proteins may also be reconstructed in artificial membrane systems to study ion channel conductance and, therefore, the "cell" employed in such assays may comprise an

artificial membrane or cell. Assays for changes in ion regulation or metabolism can be performed on cultured cells expressing endogenous normal or mutant presenilins and PS-interacting proteins. Such studies also can be performed on cells transfected with vectors capable of expressing one of the presenilins or PS-interacting proteins, or functional domains of one of the presenilins or PS-interacting proteins, in normal or mutant form. In addition, to enhance the signal measured in such assays, cells may be co-transfected with genes encoding ion channel proteins. For example, Xenopus oocytes or rat kidney (HEK293) cells may be co-transfected with sequences encoding rat brain Na^+ β1 subunits, rabbit skeletal muscle Ca^{2+} β1 subunits, or rat heart K^+ β1 subunits. Changes in presenilin or PS-interacting protein-mediated ion channel activity can be measured by, for example, two-microelectrode voltage-clamp recordings in oocytes, by whole-cell patch-clamp recordings in HEK293 cells. or by equivalent means.

C. Apoptosis or Cell Death

In another series of embodiments, compounds may be screened for their ability to modulate the activity of the presenilins or PS-interacting proteins based upon their effects on presenilin or PS-interacting protein-related apoptosis or cell death. Thus, for example, baseline rates of apoptosis or cell death may be established for cells in culture, or the baseline degree of neuronal loss at a particular age may be established post-mortem for animal models or human subjects, and the ability of a candidate compound to suppress or inhibit apoptosis or cell death may be measured. Cell death may be measured by standard microscopic techniques (e.g., light microscopy) or apoptosis may be measured more specifically by characteristic nuclear morphologies or DNA fragmentation patterns which create nucleosomal ladders (see, e.g., Gavrieli et al., 1992; Jacobson et al., 1993; Vito et al., 1996). TUNEL may also be employed to evaluate cell death in brain (see, e.g., Lassmann et al., 1995). In preferred embodiments, compounds are screened for their ability to suppress or inhibit neuronal loss in the transgenic animal models of the invention. Transgenic mice bearing, for example, a mutant human, mutant mouse, or humanized mutant presenilin or PS-interacting protein gene may be employed to identify or evaluate compounds

which may delay or arrest the neurodegeneration associated with Alzheimer's Disease. A similar transgenic mouse model, bearing a mutant APP gene, has recently been reported by Games et al. (1995).

D. Aβ Peptide Production

In another series of embodiments, compounds may be screened for their ability to modulate presenilin or PS-interacting protein-related changes in APP processing. The Aβ peptide is produced in several isoforms resulting from differences in APP processing. The Aβ peptide is a 39 to 43 amino acid derivative of βAPP which is progressively deposited in diffuse and senile plaques and in blood vessels of subjects with AD. In human brain, Aβ peptides are heterogeneous at both the N- and C-termini. Several observations, however, suggest that both the full length and N-terminal truncated forms of the long-tailed Aβ peptides ending at residue 42 or 43 (i.e., Aβ1-42/43 and Aβx-42/43) have a more important role in AD than do peptides ending at residue 40. Thus, Aβ1-42/43 and Aβx-42/43 are an early and prominent feature of both senile plaques and diffuse plaques, while peptides ending at residue 40 (i.e., Aβ1-40 and Aβx-40) are predominantly associated with a subset of mature plaques and with amyloidotic blood vessels (see, e.g., Iwatsubo et al., 1995; Gravina et al., 1995; Tamaoka et al., 1995; Podlisny et al. 1995). Furthermore, the long-tailed isoforms have a greater propensity to fibril formation, and are thought to be more neurotoxic than Aβ1-40 peptides (Pike et al., 1993; Hilbich et al., 1991). Finally, missense mutations at codon 717 of the βAPP gene are associated with early onset FAD, and result in overproduction of long-tailed Aβ in the brain of affected mutation carriers, in peripheral cells and plasma of both affected and presymptomatic carriers, and in cell lines transfected with βAPP_{717} mutant cDNAs (Tamaoka et al., 1994; Suzuki et al., 1994).

Thus, in one series of embodiments, the present invention provides methods for screening candidate compounds for their ability to block or inhibit the increased production of long isoforms of the Aβ peptides in cells or transgenic animals expressing a normal or mutant presenilin gene and/or a normal or mutant PS-interacting protein gene. In particular, the present invention provides such methods in

which cultured mammalian cells, such as brain cells or fibroblasts, have been transformed according to the methods disclosed herein, or in which transgenic animals, such as rodents or non-human primates, have been produced by the methods disclosed herein, to express relatively high levels of a normal or mutant presenilin or PS-interacting protein. Optionally, such cells or transgenic animals may also be transformed so as to express a normal or mutant form of the βAPP protein at relatively high levels.

In this series of embodiments, the candidate compound is administered to the cell line or transgenic animals (e.g., by addition to the media of cells in culture; or by oral or parenteral administration to an animal) and, after an appropriate period (e.g., 0-72 hours for cells in culture, days or months for animal models), a biological sample is collected (e.g., cell culture supernatant or cell lysate from cells in culture; tissue homogenate or plasma from an animal) and tested for the level of the long isoforms of the Aβ peptides. The levels of the peptides may be determined in an absolute sense (e.g., nMol/ml) or in a relative sense (e.g., ratio of long to short Aβ isoforms). The Aβ isoforms may be detected by any means known in the art (e.g., electrophoretic separation and sequencing) but, preferably, antibodies which are specific to the long isoform are employed to determine the absolute or relative levels of the Aβ1-42/43 or Aβx-42/43 peptides. Candidate pharmaceuticals or therapies which reduce the absolute or relative levels of these long Aβ isoforms, particularly in the transgenic animal models of the invention, are likely to have therapeutic utility in the treatment of Alzheimer's Disease, or other disorders caused by mutations in the presenilins or PS-interacting proteins, or by other aberrations in APP metabolism.

E. Phosphorylation of Microtubule Associated Proteins

In another series of embodiments, candidate compounds may be screened for their ability to modulate presenilin or PS-interacting protein activity by assessing the effect of the compound on levels of phosphorylation of microtubule associated proteins (MAPs) such as tau. The abnormal phosphorylation of tau and other MAPs in the brains of victims of Alzheimer's Disease is well known in the art. Thus, compounds which prevent or inhibit the abnormal phosphorylation of MAPs may

have utility in treating presenilin or PS-interacting protein-associated diseases such as AD. As above, cells from normal or mutant animals or subjects, or the transformed cell lines and animal models of the invention may be employed. Preferred assays will employ cell lines or animal models transformed with a mutant human or humanized mutant presenilin or PS-interacting protein gene. The baseline phosphorylation state of MAPs in these cells may be established and then candidate compounds may be tested for their ability to prevent, inhibit or counteract the hyperphosphorylation associated with mutants. The phosphorylation state of the MAPs may be determined by any standard method known in the art but, preferably, antibodies which bind selectively to phosphorylated or unphosphorylated epitopes are employed. Such antibodies to phosphorylation epitopes of the tau protein are known in the art (e.g., ALZ50).

10. Screening and Diagnostics for Alzheimer's Disease

A. General Diagnostic Methods

The PS-interacting genes and gene products, as well as the PS-interacting protein derived probes, primers and antibodies, disclosed or otherwise enabled herein, are useful in the screening for carriers of alleles associated with Alzheimer's Disease, for diagnosis of victims of Alzheimer's Disease, and for the screening and diagnosis of related presenile and senile dementias, psychiatric diseases such as schizophrenia and depression, and neurologic diseases such as stroke and cerebral hemorrhage, all of which are seen to a greater or lesser extent in symptomatic human subjects bearing mutations in the PS1 or PS2 genes or in the APP gene. Individuals at risk for Alzheimer's Disease, such as those with AD present in the family pedigree, or individuals not previously known to be at risk, may be routinely screened using probes to detect the presence of a mutant PS-interacting protein gene or protein by a variety of techniques. Diagnosis of inherited cases of these diseases can be accomplished by methods based upon the nucleic acids (including genomic and mRNA/cDNA sequences), proteins, and/or antibodies disclosed and enabled herein, including functional assays designed to detect failure or augmentation of the normal

presenilin or PS-interacting protein activity and/or the presence of specific new activities conferred by mutant PS-interacting proteins. Preferably, the methods and products are based upon the human nucleic acids, proteins or antibodies, as disclosed or otherwise enabled herein. As will be obvious to one of ordinary skill in the art, however, the significant evolutionary conservation of large portions of nucleotide and amino acid sequences, even in species as diverse as humans, mice, C. elegans, and Drosophila, allow the skilled artisan to make use of non-human homologues of the PS-interacting proteins to produce useful nucleic acids, proteins and antibodies, even for applications directed toward human or other animal subjects. Thus, for brevity of exposition, but without limiting the scope of the invention, the following description will focus upon uses of the human homologues of PS-interacting proteins and genes. It will be understood, however, that homologous sequences from other species will be equivalent for many purposes.

As will be appreciated by one of ordinary skill in the art, the choice of diagnostic methods of the present invention will be influenced by the nature of the available biological samples to be tested and the nature of the information required. Alzheimer's Disease is, of course, primarily a disease of the brain, but brain biopsies are invasive and expensive procedures, particularly for routine screening. Other tissues which express the presenilins or PS-interacting proteins at significant levels may, therefore, be preferred as sources for samples.

B. Protein Based Screens and Diagnostics

When a diagnostic assay is to be based upon PS-interacting proteins, a variety of approaches are possible. For example, diagnosis can be achieved by monitoring differences in the electrophoretic mobility of normal and mutant proteins. Such an approach will be particularly useful in identifying mutants in which charge substitutions are present, or in which insertions, deletions or substitutions have resulted in a significant change in the electrophoretic migration of the resultant protein. Alternatively, diagnosis may be based upon differences in the proteolytic cleavage patterns of normal and mutant proteins, differences in molar ratios of the

various amino acid residues, or by functional assays demonstrating altered function of the gene products.

In preferred embodiments, protein-based diagnostics will employ differences in the ability of antibodies to bind to normal and mutant PS-interacting proteins. Such diagnostic tests may employ antibodies which bind to the normal proteins but not to mutant proteins, or vice versa. In particular, an assay in which a plurality of monoclonal antibodies, each capable of binding to a mutant epitope, may be employed. The levels of anti-mutant antibody binding in a sample obtained from a test subject (visualized by, for example, radiolabelling, ELISA or chemiluminescence) may be compared to the levels of binding to a control sample. Alternatively, antibodies which bind to normal but not mutant proteins may be employed, and decreases in the level of antibody binding may be used to distinguish homozygous normal individuals from mutant heterozygotes or homozygotes. Such antibody diagnostics may be used for in situ immunohistochemistry using biopsy samples of CNS tissues obtained antemortem or postmortem, including neuropathological structures associated with these diseases such as neurofibrillary tangles and amyloid plaques, or may be used with fluid samples such a cerebrospinal fluid or with peripheral tissues such as white blood cells.

C. Nucleic Acid Based Screens and Diagnostics

When the diagnostic assay is to be based upon nucleic acids from a sample, the assay may be based upon mRNA, cDNA or genomic DNA. When mRNA is used from a sample, there are considerations with respect to source tissues and the possibility of alternative splicing. That is, there may be little or no expression of transcripts unless appropriate tissue sources are chosen or available, and alternative splicing may result in the loss of some information or difficulty in interpretation. Whether mRNA, cDNA or genomic DNA is assayed, standard methods well known in the art may be used to detect the presence of a particular sequence either in situ or in vitro (see, e.g., Sambrook et al., (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press, Cold Spring Harbor, NY). As a general matter, however, any tissue with nucleated cells may be examined.

Genomic DNA used for the diagnosis may be obtained from body cells, such as those present in the blood, tissue biopsy, surgical specimen, or autopsy material. The DNA may be isolated and used directly for detection of a specific sequence or may be amplified by the polymerase chain reaction (PCR) prior to analysis. Similarly, RNA or cDNA may also be used, with or without PCR amplification. To detect a specific nucleic acid sequence, direct nucleotide sequencing, hybridization using specific oligonucleotides, restriction enzyme digest and mapping, PCR mapping, RNase protection, chemical mismatch cleavage, ligase-mediated detection, and various other methods may be employed. Oligonucleotides specific to particular sequences can be chemically synthesized and labeled radioactively or non-radioactively (e.g., biotin tags, ethidium bromide), and hybridized to individual samples immobilized on membranes or other solid-supports (e.g., by dot-blot or transfer from gels after electrophoresis), or in solution. The presence or absence of the target sequences may then be visualized using methods such as autoradiography, fluorometry, or colorimetry. These procedures can be automated using redundant, short oligonucleotides of known sequence fixed in high density to silicon chips.

(1) Appropriate Probes and Primers

Whether for hybridization, RNase protection, ligase-mediated detection, PCR amplification or any other standards methods described herein and well known in the art, a variety of subsequences of the PS-interacting protein sequences disclosed or otherwise enabled herein will be useful as probes and/or primers. These sequences or subsequences will include both normal sequences and deleterious mutant sequences. In general, useful sequences will include at least 8-9, more preferably 10-50, and most preferably 18-24 consecutive nucleotides from introns, exons or intron/exon boundaries. Depending upon the target sequence, the specificity required, and future technological developments, shorter sequences may also have utility. Therefore, any PS-interacting protein derived sequence which is employed to isolate, clone, amplify, identify or otherwise manipulate a PS-interacting protein sequence may be regarded as an appropriate probe or primer. Particularly contemplated as useful will be sequences

including nucleotide positions from the PS-interacting protein genes in which disease-causing mutations are known to be present, or sequences which flank these positions.

(2) Hybridization Screening

For in situ detection of a normal or mutant PS-interacting protein-related nucleic acid sequence, a sample of tissue may be prepared by standard techniques and then contacted with one or more of the above-described probes, preferably one which is labeled to facilitate detection, and an assay for nucleic acid hybridization is conducted under stringent conditions which permit hybridization only between the probe and highly or perfectly complementary sequences. Because many mutations consist of a single nucleotide substitution, high stringency hybridization conditions may be required to distinguish normal sequences from most mutant sequences. When the PS-interacting protein genotypes of the subject's parents are known, probes may be chosen accordingly. Alternatively, probes to a variety of mutants may be employed sequentially or in combination. Because most individuals carrying mutations in the PS-interacting proteins will be heterozygous, probes to normal sequences also may be employed and homozygous normal individuals may be distinguished from mutant heterozygotes by the amount of binding (e.g., by intensity of radioactive signal). In another variation, competitive binding assays may be employed in which both normal and mutant probes are used but only one is labeled.

(3) Restriction Mapping

Sequence alterations may also create or destroy fortuitous restriction enzyme recognition sites which are revealed by the use of appropriate enzyme digestion followed by gel-blot hybridization. DNA fragments carrying the site (normal or mutant) are detected by their increase or reduction in size, or by the increase or decrease of corresponding restriction fragment numbers. Such restriction fragment length polymorphism analysis (RFLP), or restriction mapping, may be employed with genomic DNA, mRNA or cDNA. The PS-interacting protein sequences may be amplified by PCR using the above-described primers prior to restriction, in which case the lengths of the PCR products may indicate the presence or absence of particular restriction sites, and/or may be subjected to restriction after amplification.

The restriction fragments may be visualized by any convenient means (e.g., under UV light in the presence of ethidium bromide).

(4) PCR Mapping

In another series of embodiments, a single base substitution mutation may be detected based on differential PCR product length or production in PCR. Thus, primers which span mutant sites or which, preferably, have 3' termini at mutation sites, may be employed to amplify a sample of genomic DNA, mRNA or cDNA from a subject. A mismatch at a mutational site may be expected to alter the ability of the normal or mutant primers to promote the polymerase reaction and, thereby, result in product profiles which differ between normal subjects and heterozygous and/or homozygous mutants. The PCR products of the normal and mutant gene may be differentially separated and detected by standard techniques, such as polyacrylamide or agarose gel electrophoresis and visualization with labeled probes, ethidium bromide or the like. Because of possible non-specific priming or readthrough of mutation sites, as well as the fact that most carriers of mutant alleles will be heterozygous, the power of this technique may be low.

(5) Electrophoretic Mobility

Genetic testing based on DNA sequence differences also may be achieved by detection of alterations in electrophoretic mobility of DNA, mRNA or cDNA fragments in gels. Small sequence deletions and insertions, for example, can be visualized by high resolution gel electrophoresis of single or double stranded DNA, or as changes in the migration pattern of DNA heteroduplexes in non-denaturing gel electrophoresis. Mutations or polymorphisms in the PS-interacting protein genes may also be detected by methods which exploit mobility shifts due to single-stranded conformational polymorphisms (SSCP) associated with mRNA or single-stranded DNA secondary structures.

(6) Chemical Cleavage of Mismatches

Mutations in the PS-interacting protein genes may also be detected by employing the chemical cleavage of mismatch (CCM) method (see, e.g., Saleeba and Cotton, 1993, and references therein). In this technique, probes (up to ~ 1 kb) may be

mixed with a sample of genomic DNA, cDNA or mRNA obtained from a subject. The sample and probes are mixed and subjected to conditions which allow for heteroduplex formation (if any). Preferably, both the probe and sample nucleic acids are double-stranded, or the probe and sample may be PCR amplified together, to ensure creation of all possible mismatch heteroduplexes. Mismatched T residues are reactive to osmium tetroxide and mismatched C residues are reactive to hydroxylamine. Because each mismatched A will be accompanied by a mismatched T, and each mismatched G will be accompanied by a mismatched C, any nucleotide differences between the probe and sample (including small insertions or deletions) will lead to the formation of at least one reactive heteroduplex. After treatment with osmium tetroxide and/or hydroxylamine to modify any mismatch sites, the mixture is subjected to chemical cleavage at any modified mismatch sites by, for example, reaction with piperidine. The mixture may then be analyzed by standard techniques such as gel electrophoresis to detect cleavage products which would indicate mismatches between the probe and sample.

(7) Other Methods

Various other methods of detecting PS-interacting protein mutations, based upon the sequences disclosed and otherwise enabled herein, will be apparent to those of ordinary skill in the art. Any of these may be employed in accordance with the present invention. These include, but are not limited to, nuclease protection assays (S1 or ligase-mediated), ligated PCR, denaturing gradient gel electrophoresis (DGGE; see, e.g., Fischer and Lerman, 1983), restriction endonuclease fingerprinting combined with SSCP (REF-SSCP; see, e.g., Liu and Sommer, 1995), and the like.

D. Other Screens and Diagnostics

In inherited cases, as the primary event, and in non-inherited cases as a secondary event due to the disease state, abnormal processing of the presenilins, PS-interacting proteins, APP, or proteins reacting with the presenilins, PS-interacting proteins, or APP may occur. This can be detected as abnormal phosphorylation, glycosylation, glycation amidation or proteolytic cleavage products in body tissues or fluids (e.g., CSF or blood).

Diagnosis also can be made by observation of alterations in transcription, translation, and post-translational modification and processing, as well as alterations in the intracellular and extracellular trafficking of gene products in the brain and peripheral cells. Such changes will include alterations in the amount of messenger RNA and/or protein, alteration in phosphorylation state, abnormal intracellular location/distribution, abnormal extracellular distribution, etc. Such assays will include: Northern Blots (e.g., with PS-interacting protein-specific and non-specific nucleotide probes), Western blots and enzyme-linked immunosorbent assays (ELISA) (e.g., with antibodies raised specifically to a PS-interacting protein or PS-interacting functional domain, including various post-translational modification states including glycosylated and phosphorylated isoforms). These assays can be performed on peripheral tissues (e.g., blood cells, plasma, cultured or other fibroblast tissues, etc.) as well as on biopsies of CNS tissues obtained antemortem or postmortem, and upon cerebrospinal fluid. Such assays might also include in situ hybridization and immunohistochemistry (to localize messenger RNA and protein to specific subcellular compartments and/or within neuropathological structures associated with these diseases such as neurofibrillary tangles and amyloid plaques).

E. Screening and Diagnostic Kits

In accordance with the present invention, diagnostic kits are also provided which will include the reagents necessary for the above-described diagnostic screens. For example, kits may be provided which include antibodies or sets of antibodies which are specific to one or more mutant epitopes. These antibodies may, in particular, be labeled by any of the standard means which facilitate visualization of binding. Alternatively, kits may be provided in which oligonucleotide probes or PCR primers, as described above, are present for the detection and/or amplification of normal or mutant presenilin and/or PS-interacting protein nucleotide sequences. Again, such probes may be labeled for easier detection of specific hybridization. As appropriate to the various diagnostic embodiments described above, the oligonucleotide probes or antibodies in such kits may be immobilized to substrates and appropriate controls may be provided.

11. Methods of Treatment

The present invention now provides a basis for therapeutic intervention in diseases which are caused, or which may be caused, by mutations in the PS-interacting proteins. As noted above, mutations in the hPS1 and hPS2 genes have been associated with the development of early onset forms of Alzheimer's Disease and, therefore, the present invention is particularly directed to the treatment of subjects diagnosed with, or at risk of developing, Alzheimer's Disease.

Without being bound to any particular theory of the invention, the effect of the Alzheimer's Disease related mutations in the presenilins appears to be a gain of a novel function, or an acceleration of a normal function, which directly or indirectly causes aberrant processing of the Amyloid Precursor Protein (APP) into Aβ peptide, abnormal phosphorylation homeostasis, and/or abnormal apoptosis in the brain. Such a gain of function or acceleration of function model would be consistent with the adult onset of the symptoms and the dominant inheritance of Alzheimer's Disease. Nonetheless, the mechanism by which mutations in the presenilins may cause these effects remains unknown.

The present invention, by identifying a set of PS-interacting proteins, provides new therapeutic targets for intervening in the etiology of presenilin-related AD. In addition, as mutations in the presenilins may cause AD, it is likely that mutations in the PS-interacting proteins may also cause AD. The fact that the PS-interacting protein S5a is alternately processed in the brains of victims of sporadic AD, as well as in the brains of victims of presenilin-linked AD, suggests that, at the very least, this PS-interacting protein is involved in the etiology of AD independent of mutations in the presenilins. It is likely that the other PS-interacting proteins also may be involved in non-presenilin-linked AD.

Therapies to treat PS-interacting protein-associated diseases such as AD may be based upon (1) administration of normal PS-interacting proteins, (2) gene therapy with normal PS-interacting protein genes to compensate for or replace the mutant genes, (3) gene therapy based upon antisense sequences to mutant PS-

interacting protein genes or which "knock-out" the mutant genes, (4) gene therapy based upon sequences which encode a protein which blocks or corrects the deleterious effects of PS-interacting protein mutants, (5) immunotherapy based upon antibodies to normal and/or mutant PS-interacting proteins, or (6) small molecules (drugs) which alter PS-interacting protein expression, alter interactions between PS-interacting proteins and other proteins or ligands, or which otherwise block the aberrant function of mutant presenilin or PS-interacting proteins by altering the structure of the mutant proteins, by enhancing their metabolic clearance, or by inhibiting their function.

A. Protein Therapy

Treatment of Alzheimer's Disease, or other disorders resulting from PS-interacting protein mutations, may be performed by replacing the mutant protein with normal protein, by modulating the function of the mutant protein, or by providing an excess of normal protein to reduce the effect of any aberrant function of the mutant proteins.

To accomplish this, it is necessary to obtain, as described and enabled herein, large amounts of substantially pure PS-interacting protein from cultured cell systems which can express the protein. Delivery of the protein to the affected brain areas or other tissues can then be accomplished using appropriate packaging or administration systems including, for example, liposome mediated protein delivery to the target cells.

B. Gene Therapy

In one series of embodiments, gene therapy may be employed in which normal copies of a PS-interacting protein gene are introduced into patients to code successfully for normal protein in one or more different affected cell types. The gene must be delivered to those cells in a form in which it can be taken up and code for sufficient protein to provide effective function. Thus, it is preferred that the recombinant gene be operably joined to a strong promoter so as to provide a high level of expression which will compensate for, or out-compete, the mutant proteins. As noted above, the recombinant construct may contain endogenous or exogenous

regulatory elements, inducible or repressible regulatory elements, or tissue-specific regulatory elements.

In another series of embodiments, gene therapy may be employed to replace the mutant gene by homologous recombination with a recombinant construct. The recombinant construct may contain a normal copy of the targeted PS-interacting protein gene, in which case the defect is corrected in situ, or may contain a "knock-out" construct which introduces a stop codon, missense mutation, or deletion which abolished function of the mutant gene. It should be noted in this respect that such a construct may knock-out both the normal and mutant copies of the targeted gene in a heterozygous individual, but the total loss of gene function may be less deleterious to the individual than continued progression of the disease state.

In another series of embodiments, antisense gene therapy may be employed. The antisense therapy is based on the fact that sequence-specific suppression of gene expression can be achieved by intracellular hybridization between mRNA or DNA and a complementary antisense species. The formation of a hybrid duplex may then interfere with the transcription of the gene and/or the processing, transport, translation and/or stability of the target mRNA. Antisense strategies may use a variety of approaches including the administration of antisense oligonucleotides or antisense oligonucleotide analogs (e.g., analogs with phosphorothioate backbones) or transfection with antisense RNA expression vectors. Again, such vectors may include exogenous or endogenous regulatory regions, inducible or repressible regulatory elements, or tissue-specific regulatory elements.

In another series of embodiments, gene therapy may be used to introduce a recombinant construct encoding a protein or peptide which blocks or otherwise corrects the aberrant function caused by a mutant presenilin or PS-interacting protein gene. In one embodiment, the recombinant gene may encode a peptide which corresponds to a domain of a PS-interacting which has been found to abnormally interact with another cell protein or other cell ligand (e.g., a mutant presenilin). Thus, for example, if a mutant PS1 TM6→7 domain is found to interact with a PS-interacting protein but the corresponding normal TM6→7 domain does not undergo

this interaction, gene therapy may be employed to provide an excess of the mutant TM6→7 domain which may compete with the mutant presenilin protein and inhibit or block the aberrant interaction. Alternatively, the PS-interacting domain of a PS-interacting protein which interacts with a mutant, but not a normal, presenilin may be encoded and expressed by a recombinant construct in order to compete with, and thereby inhibit or block, the aberrant interaction.

Retroviral vectors can be used for somatic cell gene therapy especially because of their high efficiency of infection and stable integration and expression. A full length PS-interacting protein gene, subsequences encoding functional domains of these proteins, or any of the other therapeutic peptides described above, can be cloned into a retroviral vector and expression may be driven from its endogenous promoter, from the retroviral long terminal repeat, or from a promoter specific for the target cell type of interest (e.g., neurons). Other viral vectors which can be used include adeno-associated virus, vaccinia virus, bovine papilloma virus, or a herpes virus such as Epstein-Barr virus.

C. Immunotherapy

Immunotherapy is also possible for Alzheimer's Disease. Antibodies may be raised to a normal or mutant PS-interacting protein (or a portion thereof) and are administered to the patient to bind or block an aberrant interaction (e.g., with a mutant presenilin) and prevent its deleterious effects. Simultaneously, expression of the normal protein product could be encouraged. Alternatively, antibodies may be raised to specific complexes between mutant or wild-type PS-interacting proteins and their interaction partners.

A further approach is to stimulate endogenous antibody production to the desired antigen. Administration could be in the form of a one time immunogenic preparation or vaccine immunization. The PS-interacting protein or other antigen may be mixed with pharmaceutically acceptable carriers or excipients compatible with the protein. The immunogenic composition and vaccine may further contain auxiliary substances such as emulsifying agents or adjuvants to enhance effectiveness.

Immunogenic compositions and vaccines may be administered parenterally by injection subcutaneously or intramuscularly.

D. Small Molecule Therapeutics

As described and enabled herein, the present invention provides for a number of methods of identifying small molecules or other compounds which may be useful in the treatment of Alzheimer's Disease or other disorders caused by mutations in the presenilins or PS-interacting proteins. Thus, for example, the present invention provides for methods of identifying proteins which bind to normal or mutant PS-interacting proteins (aside from the presenilins). The invention also provides for methods of identifying small molecules which can be used to disrupt aberrant interactions between mutant presenilins and/or PS-interacting proteins and such other binding proteins or other cell components.

Examples

Example 1. Isolation of PS-interacting proteins by two-hybrid yeast system.

To identify proteins interacting with the presenilin proteins, a commercially available yeast two-hybrid kit ("Matchmaker System 2" from Clontech, Palo Alto, CA) was employed to screen a brain cDNA library for clones which interact with functional domains of the presenilins. In view of the likelihood that the TM6→7 loop domains of the presenilins are important functional domains, partial cDNA sequences encoding either residues 266-409 of the normal PS1 protein or residues 272-390 of the normal PS2 protein were ligated in-frame into the EcoRI and BamHI sites of the pAS2-1 fusion-protein expression vector (Clontech). The resultant fusion proteins contain the GAL4 DNA binding domain coupled in-frame either to the TM6→7 loop of the PS1 protein or to the TM6→7 loop of the PS2 protein. These expression plasmids were co-transformed into S. cerevisiae strain Y190 together with a library of human brain cDNAs ligated into the pACT2 yeast fusion-protein expression vector (Clontech) bearing the GAL4 activation domain using modified lithium acetate protocols of the "Matchmaker System 2" yeast two-hybrid kit (Clontech, Palo Alto, CA). Yeast clones bearing human brain cDNAs which interact

with the TM6→7 loop domain were selected for His- resistance by plating on SD minimal medium lacking histidine and for βgal+ activation by color selection. The His+ βgal+ clones were then purged of the pAS2-1 "bait" construct by culture in 10μg/ml cyclohexamide and the unknown "trapped" inserts of the human brain cDNAs encoding PS-interacting proteins were isolated by PCR and sequenced. Of 6 million initial transformants, 200 positive clones were obtained after His- selection, and 42 after βgal+ color selection, carried out in accordance with the manufacturer's protocol for selection of positive colonies. Of these 42 clones there were several independent clones representing the same genes.

To address the likelihood that mutations in the presenilins cause AD through the acquisition of a novel but toxic function (i.e., dominant gain of function mutation) which is mediated by a novel interaction between the mutant proteins and one or more other cellular proteins, the human brain cDNA library cloned into the pACT2 expression vector (Clontech) was re-screened using mutant TM6→7 loop domain sequences as described above and according to manufacturer's protocols. In particular, mutant presenilin sequences corresponding to residues 260-409 of PS1 TM6→7 loop domains bearing mutations L286V, L392V and Δ290-319 were ligated in-frame into the GAL4 DNA-binding domain of the pAS2-1 vector (Clontech) and used to screen the human brain cDNA:GAL4 activation domain library of pACT vectors (Clontech). Yeast were co-transformed, positive colonies were selected, and "trapped" sequences were recovered and sequenced as described above. In addition to some of the same sequences recovered with the normal TM6→7 loop domains, several new sequences were obtained which reflect aberrant interactions of the mutant presenilins with normal cellular proteins.

The recovered and sequenced clones corresponding to these PS-interacting proteins were compared to the public sequence databases using the BL:ASTN algorithm via the NCBI e-mail server. Descriptions of several of these clones follow:

Antisecretory Factor/ Proteasome S5a Subunit. Two overlapping clones (Y2H29 and Y2H31) were identified which correspond to a C-terminal fragment of a protein alternatively identified as Antisecretory Factor ("ASF") or the Multiubiquitin

chain binding S5a subunit of the 26S proteasome ("S5a") (Johansson et al. 1995; Ferrell et al., 1996). The complete nucleotide and amino acid sequences of the S5a subunit are available through the public databases under Accession number U51007 and are reproduced here as SEQ ID NO: 1 and SEQ ID NO: 2. The nucleotide sequences of the Y2H29 and Y2H31 clones include nucleotides 351-1330 of SEQ ID NO: 1 and amino acid residues 70-377 of SEQ ID NO: 2. Thus, residues 70-377 of the full S5a subunit include the PS-interacting domain of this protein. Residues 206-377 of S5a contain certain motifs that are important for protein-protein interactions (Ferrell et al., 1996).

The PS1-S5a subunit interaction was directly re-tested for both wild type and mutant PS1 TM6→7 loop (residues 260-409) by transforming Y187 yeast cells with the appropriate wild type or mutant (L286V, L392V or Δ290-319) cDNA ligated in-frame to the GAL4-DNA binding domain of pACT2. The Δ290-319 mutant fusion construct displayed autonomous βgal activation in the absence of any S5a "target sequence" and, therefore, could not be further analyzed. In contrast, both the L286V and L392V mutant constructs interacted specifically with the S5a construct. Quantitative assays, however, showed that these interactions were weaker than those involving the wild type $PS1_{260-409}$ sequence and that the degree of interaction was crudely correlated with the age of onset of FAD. The difference in βgal activation was not attributable to instability of the mutant $PS1_{260-409}$ construct mRNAs or fusion proteins because Western blots of lysates of transformed yeast showed equivalent quantities of mutant or wild-type fusion proteins.

Because one of the putative functions of S5a is to bind multi-ubiquitinated proteins, the PS1:S5a interaction observed in S. cerevisiae could arise either through yeast-dependent ubiquitination of the $PS1_{260-409}$ construct, or by direct interaction. The former would reflect a degradative pathway, a functional and perhaps reciprocal interaction between PS1 and S5a, or both. A direct interaction is favored by the fact that the PS1:S5a interaction is decreased rather than increased by the presence of the L286V and L392V mutations, and by the fact that neither of these mutations affect ubiquitin conjugation sites in the $PS1_{260-409}$ loop (i.e., K265, K311, K314 or K395).

To further examine this possibility, we investigated the direct interaction of recombinant His-tagged fusion proteins corresponding to full length S5a and the $PS1_{260-409}$ loop. Partially purified recombinant His-tagged $PS1_{260-409}$ loop and His-tagged S5a proteins and appropriate controls were mixed in phosphate buffered saline. The mixture was then subjected to size exclusion chromatography, and eluates were examined by SDS-PAGE and Western blotting using anti-His-tag monoclonal antibodies (Quiagen). In the crude $PS1_{260-409}$ loop preparation alone, the $PS1_{260-409}$ loop eluted from the size exclusion column as a broad peak at 35 minutes. In the crude S5a preparation alone, S5a eluted at 25 minutes. However, when the crude $PS1_{260-409}$ loop and S5a preparations were mixed, there was a significant shift in the elution of $PS1_{260-409}$ toward a higher molecular weight complex. Co-elution of S5a and $PS1_{260-409}$ in the same fraction was confirmed by SDS-PAGE and Western blotting of fractions using the anti-His-tag antibody. These results are consistent with a ubiquitin-independent and, therefore, possibly functional interaction.

GT24 and related genes with homology to p120/plakoglobin family. Five over-lapping clones (Y2H6, Y2H10b, Y2H17h2, Y2H24, and Y2H25) were obtained which interact with the normal PS1 TM6→7 loop domain and which appear to represent at least one novel gene. The Y2H24 clone was also found to interact with the mutant PS1 TM6→7 loop domains. Note that it appears that more than one member of the gene family was isolated, suggesting a family of genes interacting differentially with different presenilins. The most complete available cDNA corresponding to these clones was designated GT24 and is disclosed herein as SEQ ID NO: 3 and has been deposited with GenBank as Accession number U81004. The open reading frame suggests that GT24 is a protein of at least 1040 amino acids with a unique N-terminus, and considerable homology to several armadillo (arm) repeat proteins at its C-terminus. The predicted amino acid sequence of GT24 is disclosed herein as SEQ ID NO: 4. Thus, for example, residues 440-862 of GT24 have 32-56% identity ($p=1.2e^{-133}$) to residues 440-854 of murine p120 protein (Accession number Z17804), and residues 367-815 of GT24 have 26-42% identity (p=0.0017) to residues 245-465 of the D. melanogaster armadillo segment polarity protein (Accession

number P18824). The GT24 gene maps to chromosome 5p15 near the anonymous microsatellite marker D5S748 and the Cri-du-Chat syndrome locus. This sequence is also nearly identical to portions of two human ESTs of unknown function (i.e., nucleotides 2701-3018 of Accession number F08730 and nucleotides 2974-3348 of Accession number T18858). These clones also show lower degrees of homology with other partial cDNA and gDNA sequences (e.g., H17245, T06654, T77214, H24294, M62015, T87427 and G04019).

p0071 gene. An additional His$^-$, βgal$^+$ clone isolated in the initial screening with wild type $PS1_{266-409}$ "bait" had a similar nucleotide sequence to GT24 (target clone Y2H25; Accession number U81005), and would also be predicted to encode a peptide with C-terminal arm repeats. A longer cDNA sequence closely corresponding to the Y2H25 clone has been deposited in GenBank as human protein p0071 (Accession number X81889). The nucleotide and corresponding amino acid sequences of p0071 are reproduced herein as SEQ ID NOs: 5 and 6. Comparison of the predicted sequence of the p0071 ORF with that of GT24 confirms that they are related proteins with 47% overall amino acid sequence identity, and with 70% identity between residues 346-862 of GT24, and residues 509-1022 of p0071 (which includes residues encoded by the Y2H25 cDNA). The latter result strongly suggests that PS1 interacts with a novel class of arm repeat containing proteins. The broad ~ 4 kb hybridization signal obtained on Northern blots with the unique 5' end of GT24 could reflect either alternate splicing/polyadenylation of GT24, or the existence of additional members of this family with higher degrees of N-terminal homology to GT24 than p0071.

Rab11 gene. This clone (Y2H9), disclosed herein as SEQ ID NO: 7, was identified as interacting with the normal PS1 TM6→7 loop domain and appears to correspond to a known gene, Rab11, available through Accession numbers X56740 and X53143. Rab11 is believed to be involved in protein/vesicle trafficking in the ER/Golgi. Note the possible relationship to processing of membrane proteins such as ßAPP and Notch with resultant overproduction of toxic Aß peptides (especially neurotoxic $Aß_{1-42(43)}$ isoforms) (Scheuner, et al, 1995).

Retinoid X receptor-β gene. This clone (Y2H23b), disclosed herein as SEQ ID NO: 8, was identified as interacting with the normal PS1 TM6→7 loop domain and appears to correspond to a known gene, known variously as the retinoid X receptor-β, nuclear receptor co-regulator or MHC Class I regulatory element, and available through Accession numbers M84820, X63522 and M81766. This gene is believed to be involved in intercellular signaling, suggesting a possible relationship to the intercellular signaling function mediated by C. elegans sel12 and Notch/lin-12 (transcription activator).

Cytoplasmic chaperonin gene. This clone (Y2H27), disclosed herein as SEQ ID NO: 9, was identified as interacting with the normal PS1 TM6→7 loop domain and appears to correspond to a known gene, a cytoplasmic chaperonin containing TCP-1, available through Accession numbers U17104 and X74801.

Unknown gene (Y2H35). This clone (Y2H35), disclosed herein as SEQ ID NO: 10, was identified as interacting with the normal PS1 TM6→7 loop domain and appears to correspond to a known gene of unknown function, available through Accession number R12984, which shows conservation down through yeast.

Unknown gene (Y2H171). This clone (Y2H171), disclosed herein as SEQ ID NO: 11, was identified as interacting with the normal PS1 TM6→7 loop domain and appears to correspond to a known expressed repeat sequence available through Accession number D55326.

Unknown gene (Y2H41). This clone (Y2H41) was identified which reacts strongly with the TM6→7 loop domains of both PS1 and PS2 as well as the mutant loop domains of PS1. The sequence, disclosed as SEQ ID NO: 12, shows strong homology to an EST of unknown function (Accession number T64843).

Example 2. Isolation of presenilin binding proteins by affinity chromatography.

To identify the proteins which may be involved in the biochemical function of the presenilins, PS-interacting proteins were isolated using affinity chromatography. A GST-fusion protein containing the PS1 TM6→7 loop, prepared as described in Example 3, was used to probe human brain extracts, prepared by homogenizing brain tissue by Polytron in physiological salt solution. Non-specific

binding was eliminated by pre-clearing the brain homogenates of endogenous GST-binding components by incubation with glutathione-Sepharose beads. These GST-free homogenates were then incubated with the GST-PS fusion proteins to produce the desired complexes with functional binding proteins. These complexes were then recovered using the affinity glutathione-Sepharose beads. After extensive washing with phosphate buffered saline, the isolated collection of proteins was separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE; Tris-tricine gradient gel 4-20%). Two major bands were observed at ~14 and 20 kD in addition to several weaker bands ranging from 50 to 60 kD.

The same approach may now be used to identify proteins which have binding activity for the PS-interacting proteins and, thereby, to further elucidate the etiology of AD and to identify additional therapeutics targets for intervention in AD and related disorders.

Example 3. Eukaryotic and prokaryotic expression vector systems.

Constructs suitable for use in eukaryotic and prokaryotic expression systems have been generated using different classes of PS1 nucleotide cDNA sequence inserts. In the first class, termed full-length constructs, the entire PS1 cDNA sequence is inserted into the expression plasmid in the correct orientation, and includes both the natural 5' UTR and 3' UTR sequences as well as the entire open reading frame. The open reading frames bear a nucleotide sequence cassette which allows either the wild type open reading frame to be included in the expression system or alternatively, single or a combination of double mutations can be inserted into the open reading frame. This was accomplished by removing a restriction fragment from the wild type open reading frame using the enzymes NarI and PflmI and replacing it with a similar fragment generated by reverse transcriptase PCR and bearing the nucleotide sequence encoding either the M146L mutation or the H163R mutation. A second restriction fragment was removed from the wild type normal nucleotide sequence for the open reading frame by cleavage with the enzymes PflmI and NcoI and replaced with a restriction fragment bearing the nucleotide sequence encoding the A246E mutation, the A260V mutation, the A285V mutation, the L286V mutation, the

L392V mutation or the C410Y mutation. A third variant, bearing a combination of either the M146L or H163R mutation in tandem with one of the remaining mutations, was made by linking a NarI-PflmI fragment bearing one of the former mutations and a PflmI-NcoI fragment bearing one of the latter mutations.

The second class of cDNA inserts, termed truncated constructs, was constructed by removing the 5' UTR and part of the 3' UTR sequences from full length wild type or mutant cDNA sequences. The 5' UTR sequence was replaced with a synthetic oligonucleotide containing a KpnI restriction site (GGTAC/C) and a small sequence (GCCACC) to create a Kozak initiation site around the ATG at the beginning of the PS1 ORF. The 3' UTR was replaced with an oligonucleotide with an artificial EcoRI site at the 5' end. Mutant variants of this construct were then made by inserting the mutant sequences described above at the NarI-PflmI and PsImI-NcoI sites as described above.

For eukaryotic expression, these various cDNA constructs bearing wild type and mutant sequences, as described above, were cloned into the expression vector pZeoSV in which the SV60 promoter cassette had been removed by restriction digestion and replaced with the CMV promoter element of pcDNA3 (Invitrogen). For prokaryotic expression, constructs have been made using the glutathione S-transferase (GST) fusion vector pGEX-kg. The inserts which have been attached to the GST fusion nucleotide sequence are the same nucleotide sequences described above bearing either the normal open reading frame nucleotide sequence, or bearing a combination of single and double mutations as described above. These GST fusion constructs allow expression of the partial or full-length protein in prokaryotic cell systems as mutant or wild type GST fusion proteins, thus allowing purification of the full-length protein followed by removal of the GST fusion product by thrombin digestion. A further cDNA construct was made with the GST fusion vector, to allow the production of the amino acid sequence corresponding to the hydrophilic acidic loop domain between TM6 and TM7 of the full-length protein, either as a wild type nucleotide sequence or as a mutant sequence bearing either the A285V mutation, the L286V mutation or the L392V mutation. This was accomplished by recovering wild

type or mutant sequence from appropriate sources of RNA using a 5' oligonucleotide primer with a 5' BamHI restriction site (G/GATCC), and a 3' primer with a 5' EcoRI restriction site (G/AATTC). This allowed cloning of the appropriate mutant or wild type nucleotide sequence corresponding to the hydrophilic acidic loop domain at the BamHI and the EcoRI sites within the pGEX-KG vector.

The PS-interacting protein genes may be similarly manipulated by recombinant means for expression in prokaryotic or eukaryotic hosts. In particular, GST or other fusion proteins may be produced which will be useful in assays (e.g., yeast two-hybrid studies) for therapeutics.

Example 4. Antibody production.

Peptide antigens corresponding to portions of the PS1 protein were synthesized by solid-phase techniques and purified by reverse phase high pressure liquid chromatography. Peptides were covalently linked to keyhole limpet hemocyanin (KLH) via disulfide linkages that were made possible by the addition of a cysteine residue at the peptide C-terminus of the presenilin fragment. This additional residue does not appear normally in the protein sequence and was included only to facilitate linkage to the KLH molecule.

A total of three New Zealand white rabbits were immunized with peptide-KLH complexes for each peptide antigen in combination with Freund's adjuvant and were subsequently given booster injections at seven day intervals. Antisera were collected for each peptide and pooled and IgG precipitated with ammonium sulfate. Antibodies were then affinity purified with Sulfo-link agarose (Pierce) coupled with the appropriate peptide. This final purification is required to remove non-specific interactions of other antibodies present in either the pre- or post-immune serum.

The specificity of each antibody was confirmed by three tests. First, each detected single predominant bands of the approximate size predicted for presenilin-1 on Western blots of brain homogenate. Second, each cross-reacted with recombinant fusion proteins bearing the appropriate sequence. Third each could be specifically blocked by pre-absorption with recombinant PS1 or the immunizing peptide.

Antibodies to peptides derived from the PS-interacting proteins may be produced by similar means.

Example 5. Transgenic mice.

A series of wild type and mutant PS1 and PS2 genes were constructed for use in the preparation of transgenic mice. Mutant versions of PS1 and PS2 were generated by site-directed mutagenesis of the cloned cDNAs using standard techniques.

The cDNAs and their mutant versions were used to prepare two classes of mutant and wild type PS1 and PS2 cDNAs, as described in Example 3. The first class, referred to as "full-length" cDNAs, were prepared by removing approximately 200 bp of the 3' untranslated region immediately before the polyA site by digestion with EcoRI (PS1) or PvuII (PS2). The second class, referred to as "truncated" cDNAs, were prepared by replacing the 5' untranslated region with a ribosome binding site (Kozak consensus sequence) placed immediately 5' of the ATG start codon.

Various full length and truncated wild type and mutant PS1 and PS2 cDNAs, prepared as described above, were introduced into one or more of the following vectors and the resulting constructs were used as a source of gene for the production of transgenic mice.

The cos.TET expression vector: This vector was derived from a cosmid clone containing the Syrian hamster PrP gene. It has been described in detail by Scott et al. (1992) and Hsiao et al. (1995). PS1 and PS2 cDNAs (full length or truncated) were inserted into this vector at its SalI site. The final constructs contain 20 kb of 5' sequence flanking the inserted cDNA. This 5' flanking sequence includes the PrP gene promoter, 50 bp of a PrP gene 5' untranslated region exon, a splice donor site, a 1 kb intron, and a splice acceptor site located immediately adjacent to the SalI site into which the PS1 or PS2 cDNA was inserted. The 3' sequence flanking the inserted cDNA includes an approximately 8 kb segment of PrP 3' untranslated region including a polyadenylation signal. Digestion of this construct with NotI (PS1) or FseI (PS2) released a fragment containing a mutant or wild type PS gene under the

control of the PrP promoter. The released fragment was gel purified and injected into the pronuclei of fertilized mouse eggs using the method of Hsiao et al. (1995).

Platelet-derived growth factor receptor β-subunit constructs: PS cDNAs were also introduced between the SalI (full length PS1 cDNAs) or HindIII (truncated PS1 cDNAs, full length PS2 cDNAs, and truncated PS2 cDNAs) at the 3' end of the human platelet derived growth factor receptor β-subunit promoter and the EcoRI site at the 5' end of the SV40 polyA sequence and the entire cassette was cloned into the pZeoSV vector (Invitrogen, San Diego, CA.). Fragments released by ScaI/BamHI digestion were gel purified and injected into the pronuclei of fertilized mouse eggs using the method of Hsiao et al. (1995).

Human β-actin constructs: PS1 and PS2 cDNAs were inserted into the SalI site of pBAcGH. The construct produced by this insertion includes 3.4 kb of the human β actin 5' flanking sequence (the human β actin promoter, a spliced 78 bp human β actin 5' untranslated exon and intron) and the PS1 or PS2 insert followed by 2.2 kb of human growth hormone genomic sequence containing several introns and exons as well as a polyadenylation signal. SfiI was used to release a PS-containing fragment which was gel purified and injected into the pronuclei of fertilized mouse eggs using the method of Hsiao et al. (1995).

Phosphoglycerate kinase constructs: PS1 and PS2 cDNAs were introduced into the pkJ90 vector. The cDNAs were inserted between the KpnI site downstream of the human phosphoglycerate kinase promoter and the XbaI site upstream of the 3' untranslated region of the human phosphoglycerate kinase gene. PvuII/HindIII (PS1 cDNAs) or PvuII (PS2 cDNAs) digestion was used to release a PS-containing fragment which was then gel purified and injected into the pronuclei of fertilized mouse eggs as described above.

Analysis of Aβ in transgenic murine hippocampus: To analyze the effect of a mutant human PS1 transgene in mice, a PS1 mutation observed in conjunction with a particularly severe form of early-onset PS1-linked Alzheimer's disease was used, namely the M146L missense mutation (Sherrington et al., 1995). The animals, which were heterozygous for the PS1 mutant transgene on a mixed FVB-C57BL/6

strain background, were cross-bred with similar mice bearing the human wild-type βAPP_{695} cDNA under the same Syrian hamster PrP promoter similar to those animals recently described by Hsiao et al., 1995. These cross breedings were done because it is thought that human Aβ is more susceptible to the formation of aggregates than are murine Aβ peptides.

The progeny of these $PS1_{M146L}$ x βAPP_{WT} cross-breedings were then genotyped to identify animals that contained both the human wild-type βAPP_{695} transgene and also the mutant human $PS1_{M146L}$ transgene. These mice were aged until two to three months of age and then sacrificed, with the hippocampus and neocortex being dissected rapidly from the brain and frozen. Litter mates of these mice, which contained only the wild-type human βAPP_{695} transgene were also sacrificed, and their hippocampi and neocortices were dissected and rapidly frozen as well.

The concentration of both total Aβ peptides ($A\beta_{X-40}$ and $A\beta_{X-42(43)}$) as well as the subset of Aβ peptides ending on residues 42 or 43 (long-tailed $A\beta_{42}$ peptides) were then measured using a two-sandwich ELISA as described previously (Tamaoka et al., 1994; Suzuki et al., 1994). These results convincingly showed a small increase in total Aβ peptides in the double transgenic animals bearing wild-type human βAPP_{695} and mutant human $PS1_{M146L}$ transgenes compared to the wild-type human βAPP_{695} controls. More impressively, these measurements also showed that there was an increase in the amount of long-tailed Aβ peptides ending on residues 42 or 43 ($A\beta_{42}$). In contrast, litter mates bearing only the wild-type human βAPP_{695} transgene had $A\beta_{42}$ long-tailed peptide values which were below the limit of quantitation ("BLQ").

These observations therefore confirm that the construction of transgenic animals can recapitulate some of the biochemical features of human Alzheimer's disease (namely the overproduction of Aβ peptide and, in particular, overproduction of long-tailed isoforms of Aβ peptide). These observations thus prove that the transgenic models are in fact useful in exploring therapeutic targets relevant to the treatment and prevention of Alzheimer's disease.

Analysis of hippocampus dependent memory functions in PS1 transgenic mice: Fourteen transgenic C57BL/6 x FVB mice bearing the human $PS1_{M146V}$ mutant transgene under the PrP promoter (as described) above and 12 wild type litter mates aged 2.5-3 months of age (both groups were balanced for age, weight, and sex) were investigated for behavioral differences attributable to the mutant transgene. Also the qualitative observation of murine behavior in their home cages did not indicate bimodal distribution of behaviors in the sample of animals.

Experiment 1. To test for subtle differences in exploratory behavior (e.g. locomotion, scanning of the environment through rearing, and patterns of investigation of unfamiliar environment), both $PS1_{M146V}$ and wild type litter mates were tested in the open-field (Janus, et al. 1995). The results of the test revealed no significant differences between transgenics and controls in exploration of a new environment measured by mice locomotor behaviors (walking, pausing, wall leaning, rearing, grooming), ($F(1,24) = .98$, NS). Thus, differences any in behavior on the Morris water maze test (see below) cannot be attributed to differences in locomotor abilities, etc.

Experiment 2. One week after the open-field test, the $PS1_{M146V}$ mutant transgenic mice and their litter mates were trained in the Morris water maze. In this test, a mouse has to swim in a pool in order to find a submerged escape platform. The animal solves that test through learning the location of the platform using the available extra-maze spatial cues (Morris, 1990). This test was chosen because there is strong evidence that the hippocampal formation is involved in this form of learning. The hippocampus is also a major site of AD neuropathology in humans and defects in spatial learning (geographic disorientation, losing objects, wandering, etc.) are prominent early features of human AD. As a result the test is likely to detect early changes equivalent to those seen in human AD. The Morris test is conducted in three phases. In the first phase (the learning acquisition phase), the mouse has to learn the spatial position of the platform. In the second phase (the probe trial), the platform is removed from the pool and the mouse's search for the platform is recorded. In the final phase (the learning transfer phase), the platform is replaced in

a new position in the pool, and the mouse has to learn that new spatial position of the platform.

Transgenic and wild type mice did not differ in their latencies to find the platform during learning acquisition ($F(1,24) = 0.81$, NS), and both groups showed rapid learning across trials ($F(10,15) = 11.57$, $p < 0.001$). During the probe trial phase, mice from both groups searched the quadrant of the pool which originally contained the platform significantly longer than other areas of the pool which had not contained the platform ($F(3,22) = 28.9$, $p < 0.001$). However, the wild type controls showed a trend which was not quite statistically significant ($t(24) = 1.21$, $p = 0.24$) for an increased number of crossings of the exact previous position of the platform. In the learning transfer test, both groups showed the same latency of finding the new position of the platform in the initial block of trials ($t(24) = 1.11$, NS). Such long latency to find the new spatial position is expected because the mice spent most of their time searching for the platform in the old spatial position. However, in later trials in the learning transfer phase, the wild type mice showed shorter swim latencies to the new position of the platform compared to the $PS1_{M146V}$ mutant transgenics ($F(1,24)$ 2.36, $p = 0.14$). The results indicate that $PS1_{M146V}$ mutant transgenic mice were less flexible in transferring learned information to a new situation and tended to persevere in their search for the platform in the old location.

Thus, although no differences were found in the spontaneous exploration of a new environment and in the acquisition of new spatial information between the wild type and the $PS1_{M146V}$ mutant transgenic mice, the $PS1_{M146V}$ mutant transgenic mice were impaired in switching and/or adapting this knowledge in later situations.

Electrophysiological Recordings in the hippocampus of mutant transgenic mice: Five to six months old litter mate control and human $PS1_{M146V}$ mutant transgenic mice on the same C57BL/6 x FVB strain backgrounds as above were used to study long term potentiation (LTP) as an electrophysiologic correlate of learning and memory in the hippocampus. Recordings were carried out on 400 μm thick hippocampal slices according to conventional techniques. Briefly, brains were removed and transverse sections containing hippocampi were obtained within 1 min.

after mice were decapitated under halothane anesthesia. Slices were kept at room temperature in oxygenated artificial cerebrospinal fluid for one hour prior to recording. One slice at a time was transferred to the recording chamber, where they were maintained at 32 °C in an interface between oxygenated artificial cerebrospinal fluid and humidified air. Slices were then allowed to equilibrate in the recording chamber for another hour.

Extracellular field recordings were carried out in the CA1 subfield of the hippocampus at the Schaeffer collateral-pyramidal cell synapse. Synaptic responses were induced by the stimulation of Schaeffer collaterals at a frequency of 0.03 Hz and an intensity of 30-50 % of maximal response. Tetani to evoke long-term potentiation consisted of 5 trains of 100 Hz stimulation lasting for 200 ms at an intertrain interval of 10 seconds. Field potentials were recorded using an Axopatch 200B amplifier (Axon Instrument). Glass pipettes were fabricated from borosilicate glass with an outer diameter of 1.5 mm, and pulled with a two step Narishige puller. Data were acquired on a 486-IBM compatible computer using PCLAMP6 software (Axon Instrument).

To test for any abnormality in presynaptic function, we investigated the differences in paired-pulse facilitation, which is an example of use-dependent increase in synaptic efficacy and is considered to be presynaptic in origin. In hippocampus, when two stimuli are delivered to the Schaeffer collaterals in rapid succession, paired-pulse facilitation manifests itself as an enhanced dendritic response to the second stimulus as the interstimulus interval gets shorter. In three pairs of wild-type/transgenic mice, we did not observe any difference in the paired-pulse facilitation over an interstimulus interval range of 20 ms to 1 sec. These data suggest that in $PS1_{M146V}$ mutant transgenic mice, the excitability of Schaeffer collateral fibers and neurotransmitter release are likely to be normal.

Tetanic stimulation induced a long-lasting increase in the synaptic strength in both control (n = 3) and $PS1_{M146V}$ mutant transgenic mice (n = 2). In slices obtained from the $PS1_{M146V}$ mutant transgenic mice, long-lasting increase in the synaptic strength was 30 % more than that obtained from control mice.

Although preferred embodiments of the invention have been described herein in detail, it will be understood by those skilled in the art that variations may be made thereto without departing from the spirit of the invention or the scope of the appended claims.

References

Borguski et al. (1993) Nature 366:643-654.

Bueler et al. (1996) Nature 356:577-582.

Chartier-Harlin et al. (1991) Nature 353:844-846.

Ferrell et al. (1996) FEBS Lett. 381:143-148.

Fenteany et al. (1995) Science 268:726-731.

Fischer and Lerman (1983) Proc. Natl. Acad. Sci. (USA) 80:1579-1583.

Games et al. (1995) Nature 373:523-527.

Gavrieli et al. (1992) J. Cell Biol. 119:493-501.

Goate et al. (1991) Nature 349:704-706.

Gravina et al. (1995) J. Biol. Chem. 270:7013-7016.

Greco et al. (1996) Genes & Dev. 10:313-324.

Hilbich et al. (1991) J. Mol. Biol. 218:149-163.

Hogan et al. (1986) Manipulating the Mouse Embryo, Cold Spring Harbor Laboratory Press,

Cold Spring Harbor, New York.

Hsiao et al. (1995) Neuron. (in press).

Iwatsubo et al. (1995) Ann. Neurol. 37:294-299.

Jacobson et al. (1993) Nature 361:365.

Janus et al. (1995). Neurobiology of Learning and Memory, 64:58-67.

Johansson et al. (1995) J.Biol.Chem. 270:20615-20620.

Kahn (1995) Science 270:369-370.

Karlinsky et al. (1992) Neurology 42:1445-1453.

Katzman (1986) N. Eng. J. Med. 314:964-973.

Kikuchi et al. (1992) Oncogene 7:289-293.

Klafki et al. (1996) J. Biol. Chem. 271:28655-28659.

Kosik and Greenberg (1994) Alzheimer Disease. New York, Raven Press. 335-344.

Krause and Aaronson (1991) Methods in Enzymology, 200:546-556.

Lamb et al. (1993) Nature Genetics 5:22-29

Lassmann et al. (1995) Acta Neuropathol. 89:35.

Levy-Lehad et al. (1995) Science 269:970-973.

Liu and Sommer (1995) Biotechniques 18:470-477.

Morris (1990) Cold Spring Harbor Symposia on Quantitative Biology, 55:161-173.

Mullan et al. (1992) Nature Genetics 1:345-347.

Murrell et al. (1991) Science 254:97-99.

Nowak (1995) Science 270:368-371.

Okamoto, et al. (1995) J. Biol. Chem. 270:4205-4208.

Phizicky and Fields (1994) Microbiol. Reviews 59:94-123.

Pike et al. (1993) J. Neurosci. 13:1676-1678.

Podlisny et al. (1995) J. Biol. Chem. 270:9564-9570.

Rock et al. (1994) Cell 78:761-771.

Rogaev et al. (1995) Nature 376:775-778.

Saleeba and Cotton (1993) Methods in Enzymology 217:286-295.

Scheuner et al. (1995) Soc. Neurosci. Abstr. 21:1500.

Scott et al. (1992) Protein Sci. 1:986-997.

Sherrington et al. (1995) Nature 375:754-760.

Suzuki et al. (1994) Science 264:1336-1340.

Takada et al. (1994) Genes & Dev. 8:174-189.

Tamaoka et al. (1994) J. Biol. Chem. 269:32721-32724.

Tamaoka et al. (1995) Brain Res. 679:151-156.

Vito et al. (1996) Science 271:521-525.

Wolozin et al. (1996) Science 274:1710-1713.

Wong et al. (1996) Neuroscience 22:728.

Yamatsuji et al. (1996) Science 272:1349-1352.

Zheng et al. (1995) Cell 81:525-531

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: ST. GEORGE-HYSLOP, PETER H.
ROMMENS, JOHANNA M.
FRASER, PAUL E.

(ii) TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS RELATED TO ALZHEIMER'S DISEASE, AND USES THEREFOR

(iii) NUMBER OF SEQUENCES: 12

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Sim & McBurney
(B) STREET: 330 University Avenue, 6TH Floor
(C) CITY: Toronto
(D) STATE: Ontario
(E) COUNTRY: Canada
(F) ZIP: M5G 1R7

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 27-JAN-1997
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/592,541
(B) FILING DATE: 26-JAN-1996

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/021,673
(B) FILING DATE: 05-JUL-1996

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/021,700
(B) FILING DATE: 12-JUL-1996

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 60/029,895
(B) FILING DATE: 08-NOV-1996

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: DOCKET# CAN-006PR(d)
(B) FILING DATE: 02-JAN-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: RAE, Patricia A.

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (416) 595-1155
(B) TELEFAX: (416) 595-1163

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1330 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 145..1275
(D) OTHER INFORMATION: /product= "S5a"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
AATTCCCAAA TGACCTTTTA TTTCATACAG AGATACAAAG GCAACTATGT GCAGCAACAA      60

TCTGATGGGC AGTCCAAACT CTTGGGAGGA AGTAAATTCA TGGTAAATGT CATGATGGCG     120

GTCGGGAGGG AGGAAGGTGG CAAG ATG GTG TTG GAA AGC ACT ATG GTG TGT         171
                           Met Val Leu Glu Ser Thr Met Val Cys
                             1               5

GTG GAC AAC AGT GAG TAT ATG CGG AAT GGA GAC TTC TTA CCC ACC AGG        219
Val Asp Asn Ser Glu Tyr Met Arg Asn Gly Asp Phe Leu Pro Thr Arg
 10                  15                  20                  25

CTG CAG GCC CAG CAG GAT GCT GTC AAC ATA GTT TGT CAT TCA AAG ACC        267
Leu Gln Ala Gln Gln Asp Ala Val Asn Ile Val Cys His Ser Lys Thr
                 30                  35                  40
```

```
CGC AGC AAC CCT GAG AAC AAC GTG GGC CTT ATC ACA CTG GCT AAT GAC          411
Arg Ser Asn Pro Glu Asn Asn Val Gly Leu Ile Thr Leu Ala Asn Asp
                45                  50                  55

TGT GAA GTG CTG ACC ACA CTC ACC CCA GAC ACT GGC CGT ATC CTG TCC          459
Cys Glu Val Leu Thr Thr Leu Thr Pro Asp Thr Gly Arg Ile Leu Ser
            60                  65                  70

AAG CTA CAT ACT GTC CAA CCC AAG GGC AAG ATC ACC TTC TGC ACG GGC          411
Lys Leu His Thr Val Gln Pro Lys Gly Lys Ile Thr Phe Cys Thr Gly
        75                  80                  85

ATC CGC GTG GCC CAT CTG GCT CTG AAG CAC CGA CAA GGC AAG AAT CAC          459
Ile Arg Val Ala His Leu Ala Leu Lys His Arg Gln Gly Lys Asn His
    90                  95                  100                 105

AAG ATG CGC ATC ATT GCC TTT GTG GGA AGC CCA GTG GAG GAC AAT GAG          507
Lys Met Arg Ile Ile Ala Phe Val Gly Ser Pro Val Glu Asp Asn Glu
                110                 115                 120

AAG GAT CTG GTG AAA CTG GCT AAA CGC CTC AAG AAG GAG AAA GTA AAT          555
Lys Asp Leu Val Lys Leu Ala Lys Arg Leu Lys Lys Glu Lys Val Asn
            125                 130                 135

GTT GAC ATT ATC AAT TTT GGG GAA GAG GAG GTG AAC ACA GAA AAG CTG          603
Val Asp Ile Ile Asn Phe Gly Glu Glu Glu Val Asn Thr Glu Lys Leu
        140                 145                 150

ACA GCC TTT GTA AAC ACG TTG AAT GGC AAA GAT GGA ACC GGT TCT CAT          651
Thr Ala Phe Val Asn Thr Leu Asn Gly Lys Asp Gly Thr Gly Ser His
    155                 160                 165

CTG GTG ACA GTG CCT CCT GGG CCC AGT TTG GCT GAT GCT CTC ATC AGT          699
Leu Val Thr Val Pro Pro Gly Pro Ser Leu Ala Asp Ala Leu Ile Ser
170                 175                 180                 185

TCT CCG ATT TTG GCT GGT GAA GGT GGT GCC ATG CTG GGT CTT GGT GCC          747
Ser Pro Ile Leu Ala Gly Glu Gly Gly Ala Met Leu Gly Leu Gly Ala
                190                 195                 200

AGT GAC TTT GAA TTT GGA GTA GAT CCC AGT GCT GAT CCT GAG CTG GCC          795
Ser Asp Phe Glu Phe Gly Val Asp Pro Ser Ala Asp Pro Glu Leu Ala
            205                 210                 215

TTG GCC CTT CGT GTA TCT ATG GAA GAG CAG CGG CAG CGG CAG GAG GAG          843
Leu Ala Leu Arg Val Ser Met Glu Glu Gln Arg Gln Arg Gln Glu Glu
        220                 225                 230

GAG GCC CGG CGG GCA GCT GCA GCT TCT GCT GCT GAG GCC GGG ATT GCT          891
Glu Ala Arg Arg Ala Ala Ala Ala Ser Ala Ala Glu Ala Gly Ile Ala
    235                 240                 245

ACG ACT GGG ACT GAA GAC TCA GAC GAT GCC CTG CTG AAG ATG ACC ATC          939
Thr Thr Gly Thr Glu Asp Ser Asp Asp Ala Leu Leu Lys Met Thr Ile
250                 255                 260                 265

AGC CAG CAA GAG TTT GGC CGC ACT GGG CTT CCT GAC CTA AGC AGT ATG          987
Ser Gln Gln Glu Phe Gly Arg Thr Gly Leu Pro Asp Leu Ser Ser Met
                270                 275                 280

ACT GAG GAA GAG CAG ATT GCT TAT GCC ATG CAG ATG TCC CTG CAG GGA         1035
Thr Glu Glu Glu Gln Ile Ala Tyr Ala Met Gln Met Ser Leu Gln Gly
            285                 290                 295

GCA GAG TTT GGC CAG GCG GAA TCA GCA GAC ATT GAT GCC AGC TCA GCT         1083
Ala Glu Phe Gly Gln Ala Glu Ser Ala Asp Ile Asp Ala Ser Ser Ala
        300                 305                 310

ATG GAC ACA TCC GAG CCA GCC AAG GAG GAG GAT GAT TAC GAC GTG ATG         1131
Met Asp Thr Ser Glu Pro Ala Lys Glu Glu Asp Asp Tyr Asp Val Met
    315                 320                 325

CAG GAC CCC GAG TTC CTT CAG AGT GTC CTA GAG AAC CTC CCA GGT GTG         1179
Gln Asp Pro Glu Phe Leu Gln Ser Val Leu Glu Asn Leu Pro Gly Val
330                 335                 340                 345

GAT CCC AAC AAT GAA GCC ATT CGA AAT GCT ATG GGC TCC CTG GCC TCC         1227
Asp Pro Asn Asn Glu Ala Ile Arg Asn Ala Met Gly Ser Leu Ala Ser
                350                 355                 360

CAG GCC ACC AAG GAC GGC AAG AAG GAC AAG AAG GAG GAA GAC AAG AAG         1275
Gln Ala Thr Lys Asp Gly Lys Lys Asp Lys Lys Glu Glu Asp Lys Lys
            365                 370                 375

TGAGACTGGA GGGAAAGGGT AGCTGAGTCT GCTTAGGGAC TGCATGGGGG AATTC            1330
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 377 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Val Leu Glu Ser Thr Met Val Cys Val Asp Asn Ser Glu Tyr Met
1 5 10 15

Arg Asn Gly Asp Phe Leu Pro Thr Arg Leu Gln Ala Gln Gln Asp Ala
20 25 30

Val Asn Ile Val Cys His Ser Lys Thr Arg Ser Asn Pro Glu Asn Asn
35 40 45

Val Gly Leu Ile Thr Leu Ala Asn Asp Cys Glu Val Leu Thr Thr Leu
50 55 60

Thr Pro Asp Thr Gly Arg Ile Leu Ser Lys Leu His Thr Val Gln Pro
65 70 75 80

Lys Gly Lys Ile Thr Phe Cys Thr Gly Ile Arg Val Ala His Leu Ala
85 90 95

Leu Lys His Arg Gln Gly Lys Asn His Lys Met Arg Ile Ile Ala Phe
100 105 110

Val Gly Ser Pro Val Glu Asp Asn Glu Lys Asp Leu Val Lys Leu Ala
115 120 125

Lys Arg Leu Lys Lys Glu Lys Val Asn Val Asp Ile Ile Asn Phe Gly
130 135 140

Glu Glu Glu Val Asn Thr Glu Lys Leu Thr Ala Phe Val Asn Thr Leu
145 150 155 160

Asn Gly Lys Asp Gly Thr Gly Ser His Leu Val Thr Val Pro Pro Gly
165 170 175

Pro Ser Leu Ala Asp Ala Leu Ile Ser Ser Pro Ile Leu Ala Gly Glu
180 185 190

Gly Gly Ala Met Leu Gly Leu Gly Ala Ser Asp Phe Glu Phe Gly Val
195 200 205

Asp Pro Ser Ala Asp Pro Glu Leu Ala Leu Ala Leu Arg Val Ser Met
210 215 220

Glu Glu Gln Arg Gln Arg Gln Glu Glu Glu Ala Arg Arg Ala Ala Ala
225 230 235 240

Ala Ser Ala Ala Glu Ala Gly Ile Ala Thr Thr Gly Thr Glu Asp Ser
245 250 255

Asp Asp Ala Leu Leu Lys Met Thr Ile Ser Gln Gln Glu Phe Gly Arg
260 265 270

Thr Gly Leu Pro Asp Leu Ser Ser Met Thr Glu Glu Glu Gln Ile Ala
275 280 285

Tyr Ala Met Gln Met Ser Leu Gln Gly Ala Glu Phe Gly Gln Ala Glu
290 295 300

Ser Ala Asp Ile Asp Ala Ser Ser Ala Met Asp Thr Ser Glu Pro Ala
305 310 315 320

Lys Glu Glu Asp Asp Tyr Asp Val Met Gln Asp Pro Glu Phe Leu Gln
325 330 335

Ser Val Leu Glu Asn Leu Pro Gly Val Asp Pro Asn Asn Glu Ala Ile
340 345 350

Arg Asn Ala Met Gly Ser Leu Ala Ser Gln Ala Thr Lys Asp Gly Lys
355 360 365

Lys Asp Lys Lys Glu Glu Asp Lys Lys
370 375

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3841 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 2..3121

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..3841
(D) OTHER INFORMATION: /note= "GT24"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
T TCA CAG CTC CCG GCC CGA GGC ACA CAA GCC CGA GST ACG GGC CAG                 46
  Ser Gln Leu Pro Ala Arg Gly Thr Gln Ala Arg Xaa Thr Gly Gln
    1               5                  10                  15

AGC TTC AGC CAG GGC ACG ACC AGC CGC GCC GGC CAC CTG GCG GGG CCC              94
Ser Phe Ser Gln Gly Thr Thr Ser Arg Ala Gly His Leu Ala Gly Pro
                20                  25                  30

GAG CCC GCG CCG CCG CCG CCG CCG CCR CCG CGG GAG CCG TTC GCG CCC             142
Glu Pro Ala Pro Pro Pro Pro Pro Xaa Pro Arg Glu Pro Phe Ala Pro
            35                  40                  45

AGC CTG GGC AGC GCC TTC CAC CTG CCC GAC GCG CCG CCC GCC GCC GCC             190
Ser Leu Gly Ser Ala Phe His Leu Pro Asp Ala Pro Pro Ala Ala Ala
        50                  55                  60

GCC GCC GCG CTC TAC TAC TCC AKC TCC ACG CTG CCC GCG CCG CCG CGC             238
Ala Ala Ala Leu Tyr Tyr Ser Xaa Ser Thr Leu Pro Ala Pro Pro Arg
    65                  70                  75

GGG GGC TCC CCG CTG GCC GCG CCC CAG GGC GGT TCG CCC ACC AAG CTG             286
Gly Gly Ser Pro Leu Ala Ala Pro Gln Gly Gly Ser Pro Thr Lys Leu
80                  85                  90                  95

CAG CGC GGC GGC TCG GCC CCC GAG GGC GCC ACC TAC GCC GCG CCG CGC             334
Gln Arg Gly Gly Ser Ala Pro Glu Gly Ala Thr Tyr Ala Ala Pro Arg
                100                 105                 110

GGC TCC TCG CCC AAG CAG TCG CCC AGC CGC CTG GCC AAG TCC TAC AGC             382
Gly Ser Ser Pro Lys Gln Ser Pro Ser Arg Leu Ala Lys Ser Tyr Ser
            115                 120                 125

ACC AGC TCG CCC ATC AAC ATC GTC GTG TCC TCG GCC GGC CTG TCC CCG             430
Thr Ser Ser Pro Ile Asn Ile Val Val Ser Ser Ala Gly Leu Ser Pro
        130                 135                 140

ATC CGC GTG ACC TCG CCC CCC ACC GTG CAG TCC ACC ATC TCC TCC TCG             478
Ile Arg Val Thr Ser Pro Pro Thr Val Gln Ser Thr Ile Ser Ser Ser
    145                 150                 155

CCC ATC CAC CAG CTG AGC TCC ACC ATC GGC ACG TAC GCC ACC CTG TCG             526
Pro Ile His Gln Leu Ser Ser Thr Ile Gly Thr Tyr Ala Thr Leu Ser
160                 165                 170                 175

CCC ACC AAG CGC CTG GTC CAC GCG TCC GAG CAG TAC AGC AAG CAC TCG             574
Pro Thr Lys Arg Leu Val His Ala Ser Glu Gln Tyr Ser Lys His Ser
                180                 185                 190

CAG GAG CTG TAT GCC ACG GCC ACC CTC CAG AGG CCG GGC AGC CTG GCA             622
Gln Glu Leu Tyr Ala Thr Ala Thr Leu Gln Arg Pro Gly Ser Leu Ala
            195                 200                 205

GCT GGT TCC CGA GCC TCA TAC AGC AGC CAG CAT GGG CAC CTG GGC CCA             670
Ala Gly Ser Arg Ala Ser Tyr Ser Ser Gln His Gly His Leu Gly Pro
        210                 215                 220

GAG TTG CGG GCC CTG CAG TCC CCA GAA CAC CAC ATA GAT CCC ATC TAT             718
Glu Leu Arg Ala Leu Gln Ser Pro Glu His His Ile Asp Pro Ile Tyr
    225                 230                 235

GAA GAC CGC GTC TAT CAG AAG CCC CCT ATG AGG AGT CTC AGC CAG AGC             766
Glu Asp Arg Val Tyr Gln Lys Pro Pro Met Arg Ser Leu Ser Gln Ser
240                 245                 250                 255

CAG GGG GAC CCT CTG CCG CCA GCA CAC ACC GGC ACC TAC CGC ACG AGC             814
Gln Gly Asp Pro Leu Pro Pro Ala His Thr Gly Thr Tyr Arg Thr Ser
                260                 265                 270

ACA GCC CCA TCT TCC CCT GGT GTC GAC TCC GTC CCC TTG CAG CGC ACA             862
Thr Ala Pro Ser Ser Pro Gly Val Asp Ser Val Pro Leu Gln Arg Thr
            275                 280                 285

GGC AGC CAG CAC GGC CCA CAG AAT GCC GCC GCG GCC ACC TTC CAG AGG             910
Gly Ser Gln His Gly Pro Gln Asn Ala Ala Ala Ala Thr Phe Gln Arg
        290                 295                 300

GCC AGC TAT GCC GCC GGC CCA GCC TCC AAT TAC GCG GAC CCC TAC CGA             958
Ala Ser Tyr Ala Ala Gly Pro Ala Ser Asn Tyr Ala Asp Pro Tyr Arg
    305                 310                 315

CAG CTG CAG TAT TGT CCC TCT GTT GAG TCT CCA TAC AGC AAA TCC GGC            1006
Gln Leu Gln Tyr Cys Pro Ser Val Glu Ser Pro Tyr Ser Lys Ser Gly
320                 325                 330                 335

CCT GCT CTC CCG CCT GAA GGC ACC TTG GCC AGG TCC CCG TCC ATT GAT            1054
Pro Ala Leu Pro Pro Glu Gly Thr Leu Ala Arg Ser Pro Ser Ile Asp
                340                 345                 350

AGC ATT CAG AAA GAT CCC AGA GAA TTT GGA TGG AGA GAC CCG GAA CTG            1102
Ser Ile Gln Lys Asp Pro Arg Glu Phe Gly Trp Arg Asp Pro Glu Leu
            355                 360                 365

CCG GAA GTG ATT CAG ATG TTG CAG CAC CAG TTT CCC TCG GTC CAG TCT            1150
```

```
Pro Glu Val Ile Gln Met Leu Gln His Gln Phe Pro Ser Val Gln Ser
           370                 375                 380

AAC GCG GCA GCC TAC TTG CAA CAC CTC TGT TTT GGA GAC AAC AAA ATT      1198
Asn Ala Ala Ala Tyr Leu Gln His Leu Cys Phe Gly Asp Asn Lys Ile
385                 390                 395

AAA GCC GAG ATA AGG AGA CAA GGA GGC ATC CAG CTC CTG GTG GAC CTG      1246
Lys Ala Glu Ile Arg Arg Gln Gly Gly Ile Gln Leu Leu Val Asp Leu
400                 405                 410                 415

TTG GAT CAT CGG ATG ACC GAA GTC CAC CGT AGT GCC TGT GGA GCT CTG      1294
Leu Asp His Arg Met Thr Glu Val His Arg Ser Ala Cys Gly Ala Leu
                420                 425                 430

AGA AAC CTG GTG TAT GGG AAG GCC AAC GAT GAT AAC AAA ATT GCC CTG      1342
Arg Asn Leu Val Tyr Gly Lys Ala Asn Asp Asp Asn Lys Ile Ala Leu
            435                 440                 445

AAA AAC TGT GGT GGC ATC CCA GCA CTG GTG AGG TTA CTC CGC AAG ACG      1390
Lys Asn Cys Gly Gly Ile Pro Ala Leu Val Arg Leu Leu Arg Lys Thr
        450                 455                 460

ACT GAC CTG GAG ATC CGG GAG CTG GTC ACA GGA GTC CTT TGG AAC CTC      1438
Thr Asp Leu Glu Ile Arg Glu Leu Val Thr Gly Val Leu Trp Asn Leu
465                 470                 475

TCC TCA TGC GAT GCA CTC AAA ATG CCA ATC ATC CAG GAT GCC CTA GCA      1486
Ser Ser Cys Asp Ala Leu Lys Met Pro Ile Ile Gln Asp Ala Leu Ala
480                 485                 490                 495

GTA CTG ACC AAC GCG GTG ATT ATC CCC CAC TCA GGC TGG GAA AAT TCG      1534
Val Leu Thr Asn Ala Val Ile Ile Pro His Ser Gly Trp Glu Asn Ser
                500                 505                 510

CCT CTT CAG GAT GAT CGG AAA ATA CAG CTG CAT TCA TCA CAG GTG CTG      1582
Pro Leu Gln Asp Asp Arg Lys Ile Gln Leu His Ser Ser Gln Val Leu
            515                 520                 525

CGT AAC GCC ACC GGG TGC CTA AGG AAT GTT AGT TCG GCC GGA GAG GAG      1630
Arg Asn Ala Thr Gly Cys Leu Arg Asn Val Ser Ser Ala Gly Glu Glu
        530                 535                 540

GCC CGC AGA AGG ATG AGA GAG TGT GAT GGG CTT ACG GAT GCC TTG CTG      1678
Ala Arg Arg Arg Met Arg Glu Cys Asp Gly Leu Thr Asp Ala Leu Leu
545                 550                 555

TAC GTG ATC CAG TCT GCG CTG GGG AGC AGT GAG ATC GAT AGC AAG ACC      1726
Tyr Val Ile Gln Ser Ala Leu Gly Ser Ser Glu Ile Asp Ser Lys Thr
560                 565                 570                 575

GTT GAA AAC TGT GTG TGC ATT TTA AGG AAC CTC TCG TAC CGG CTG GCG      1774
Val Glu Asn Cys Val Cys Ile Leu Arg Asn Leu Ser Tyr Arg Leu Ala
                580                 585                 590

GCA GAA ACG TCT CAG GGA CAG CAC ATG GGC ACG GAC GAG CTG GAC GGG      1822
Ala Glu Thr Ser Gln Gly Gln His Met Gly Thr Asp Glu Leu Asp Gly
            595                 600                 605

CTA CTC TGT GGC GAG GCC AAT GGC AAG GAT GCT GAG AGC TCT GGG TGC      1870
Leu Leu Cys Gly Glu Ala Asn Gly Lys Asp Ala Glu Ser Ser Gly Cys
        610                 615                 620

TGG GGC AAG AAG AAG AAG AAA AAG AAA TCC CAA GAT CAG TGG GAT GGA      1918
Trp Gly Lys Lys Lys Lys Lys Lys Lys Ser Gln Asp Gln Trp Asp Gly
625                 630                 635

GTA GGA CCT CTT CCA GAC TGT GCT GAA CCA CCA AAA GGG ATC CAG ATG      1966
Val Gly Pro Leu Pro Asp Cys Ala Glu Pro Pro Lys Gly Ile Gln Met
640                 645                 650                 655

CTG TGG CAC CCA TCA ATA GTC AAA CCC TAC CTC ACA CTG CTC TCT GAG      2014
Leu Trp His Pro Ser Ile Val Lys Pro Tyr Leu Thr Leu Leu Ser Glu
                660                 665                 670

TGC TCA AAT CCA GAC ACG CTG GAA GGG GCG GCA GGC GCC CTG CAG AAC      2062
Cys Ser Asn Pro Asp Thr Leu Glu Gly Ala Ala Gly Ala Leu Gln Asn
            675                 680                 685

TTG GCT GCA GGG AGC TGG AAG TGG TCA GTA TAT ATC CGA GCC GCT GTC      2110
Leu Ala Ala Gly Ser Trp Lys Trp Ser Val Tyr Ile Arg Ala Ala Val
        690                 695                 700

CGA AAA GAG AAA GGC CTG CCC ATC CTC GTG GAG CTG CTC CGA ATA GAC      2158
Arg Lys Glu Lys Gly Leu Pro Ile Leu Val Glu Leu Leu Arg Ile Asp
705                 710                 715

AAT GAC CGT GTG GTG TGC GCG GTG GCC ACT GCG CTG CGG AAC ATG GCC      2206
Asn Asp Arg Val Val Cys Ala Val Ala Thr Ala Leu Arg Asn Met Ala
720                 725                 730                 735

TTG GAC GTC AGA AAT AAG GAG CTC ATC GGC AAA TAC GCC ATG CGA GAC      2254
Leu Asp Val Arg Asn Lys Glu Leu Ile Gly Lys Tyr Ala Met Arg Asp
                740                 745                 750
```

```
CTA GTC CAC AGG CTT CCA GGA GGG AAC AAC AGC AAC AAC ACT GCA AGC
Leu Val His Arg Leu Pro Gly Gly Asn Asn Ser Asn Asn Thr Ala Ser
            755                 760                 765

AAG GCC ATG TCG GAT GAC ACA GTG ACA GCT GTC TGC TGC ACA CTG CAC
Lys Ala Met Ser Asp Asp Thr Val Thr Ala Val Cys Cys Thr Leu His
        770                 775                 780

GAA GTG ATT ACC AAG AAC ATG GAG AAC GCC AAG GCC TTA CGG GAT GCC
Glu Val Ile Thr Lys Asn Met Glu Asn Ala Lys Ala Leu Arg Asp Ala
    785                 790                 795

GGT GGC ATC GAG AAG TTG GTC GGC ATC TCC AAA AGC AAA GGA GAT AAA
Gly Gly Ile Glu Lys Leu Val Gly Ile Ser Lys Ser Lys Gly Asp Lys
800                 805                 810                 815

CAC TCT CCA AAA GTG GTC AAG GCT GCA TCT CAG GTC CTC AAC AGC ATG
His Ser Pro Lys Val Val Lys Ala Ala Ser Gln Val Leu Asn Ser Met
                820                 825                 830

TGG CAG TAC CGA GAT CTG AGG AGT CTC TAC AAA AAG GAT GGA TGG TCA
Trp Gln Tyr Arg Asp Leu Arg Ser Leu Tyr Lys Lys Asp Gly Trp Ser
            835                 840                 845

CAA TAC CAC TTT GTA GCC TCG TCT TCA ACC ATC GAG AGG GAC CGG CAA
Gln Tyr His Phe Val Ala Ser Ser Ser Thr Ile Glu Arg Asp Arg Gln
        850                 855                 860

AGG CCC TAC TCC TCC TCC CGC ACG CCC TCC ATC TCC CCT GTG CGC GTG
Arg Pro Tyr Ser Ser Ser Arg Thr Pro Ser Ile Ser Pro Val Arg Val
    865                 870                 875

TCT CCC AAC AAC CGC TCA GCA AGT GCC CCA GCT TCA CCT CGG GAA ATG
Ser Pro Asn Asn Arg Ser Ala Ser Ala Pro Ala Ser Pro Arg Glu Met
880                 885                 890                 895

ATC AGC CTC AAA GAA AGG AAA ACA GAC TAC GAG TGC ACC GGC AGC AAC
Ile Ser Leu Lys Glu Arg Lys Thr Asp Tyr Glu Cys Thr Gly Ser Asn
                900                 905                 910

GCC ACC TAC CAC GGA GGT AAA GGC GAA CAC ACT TCC AGG AAA GAT GCC
Ala Thr Tyr His Gly Gly Lys Gly Glu His Thr Ser Arg Lys Asp Ala
            915                 920                 925

ATG ACA GCT CAA AAC ACT GGA ATT TCA ACT TTG TAT AGG AAT TCT TAT
Met Thr Ala Gln Asn Thr Gly Ile Ser Thr Leu Tyr Arg Asn Ser Tyr
        930                 935                 940

GGT GCG CCC GCT GAA GAC ATC AAA CAC AAC CAG GTT TCA GCA CAG CCA
Gly Ala Pro Ala Glu Asp Ile Lys His Asn Gln Val Ser Ala Gln Pro
    945                 950                 955

GTC CCA CAG GAG CCC AGC AGA AAA GAT TAC GAG ACC TAC CAG CCA TTT
Val Pro Gln Glu Pro Ser Arg Lys Asp Tyr Glu Thr Tyr Gln Pro Phe
960                 965                 970                 975

CAG AAT TCC ACA AGA AAT TAC GAT GAG TCC TTC TTC GAG GAC CAG GTC
Gln Asn Ser Thr Arg Asn Tyr Asp Glu Ser Phe Phe Glu Asp Gln Val
                980                 985                 990

CAC CAT CGC CCT CCC GCC AGC GAG TAC ACC ATG CAC CTG GGT CTC AAG
His His Arg Pro Pro Ala Ser Glu Tyr Thr Met His Leu Gly Leu Lys
            995                1000                1005

TCC ACC GGC AAC TAC GTT GAC TTC TAC TCA GCT GCC CGT CCC TAC AGT
Ser Thr Gly Asn Tyr Val Asp Phe Tyr Ser Ala Ala Arg Pro Tyr Ser
        1010                1015                1020

GAA CTG AAC TAT GAA ACG AGC CAC TAC CCG GCC TCC CCC GAC TCC TGG
Glu Leu Asn Tyr Glu Thr Ser His Tyr Pro Ala Ser Pro Asp Ser Trp
    1025                1030                1035

GTG TGAGGAGCAG GCCACAGGCG CTCCGGGAAA CAGTGCATGT GCATGCATAC
Val
1040

CACAAGACAT TTCTTTCTGT TTTGGTTTTT TTCTCCTGCA AATTTAGTTT GTTAAAGCCT

GTTCCATAGG AAGGCTGTGA TAACCAGTAA GGGAAATATT AAGAGCTATT TTAGAAAGCT

AAATGAATCG CAAGTTAACT TGGAAATCAG TAGAAAGCTA AAGTGATCCT AAATATGACA

GTGGGCAGCA CCTTTCCTAG CGTGTTNTGT TAGGAGTAAC GAGAAGTGCT TTATACTGAA

CGTGGGTTGN TTGGTAGGGT GGAGNCGAGG CATTCGGGCC GGTGGGGCGT AAGGGTTATC

GTTAAGCACA AGACACAGAA TAGTTTACAC ACTGTGTGGG GGACGGCTTC TCACGCTTTG

TTTACTCTCT TCATCCGTTG TGACTCTAGG CTTCAGGTTG CATTGGGGTT CCTCTGTACA

GCAAGATGTT TCTTGCCTTT TGTTAATGCA TTGTTGTAAA GTATTTGATG TACATTACAG

ATTAAAGAAG NAAAGCGCGT TGTGTATATT ACACCAATNC CGCCGTGTTT CCTCATCTAT

GGTTCTAAAT ATTGCTTCAA TTTCNAACTT TTGAAAGATG TATGGATTTC CAGTTTTTCT
```

TTACTTTCTC CCAGTATGTT TTAACGNMMN AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1941

AAAACTCGAG 1961

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1040 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ser Gln Leu Pro Ala Arg Gly Thr Gln Ala Arg Xaa Thr Gly Gln Ser
1 5 10 15

Phe Ser Gln Gly Thr Thr Ser Arg Ala Gly His Leu Ala Gly Pro Glu
20 25 30

Pro Ala Pro Pro Pro Pro Pro Xaa Pro Arg Glu Pro Phe Ala Pro Ser
35 40 45

Leu Gly Ser Ala Phe His Leu Pro Asp Ala Pro Pro Ala Ala Ala Ala
50 55 60

Ala Ala Leu Tyr Tyr Ser Xaa Ser Thr Leu Pro Ala Pro Pro Arg Gly
65 70 75 80

Gly Ser Pro Leu Ala Ala Pro Gln Gly Gly Ser Pro Thr Lys Leu Gln
85 90 95

Arg Gly Gly Ser Ala Pro Glu Gly Ala Thr Tyr Ala Ala Pro Arg Gly
100 105 110

Ser Ser Pro Lys Gln Ser Pro Ser Arg Leu Ala Lys Ser Tyr Ser Thr
115 120 125

Ser Ser Pro Ile Asn Ile Val Ser Ser Ala Gly Leu Ser Pro Ile
130 135 140

Arg Val Thr Ser Pro Pro Thr Val Gln Ser Thr Ile Ser Ser Ser Pro
145 150 155 160

Ile His Gln Leu Ser Ser Thr Ile Gly Thr Tyr Ala Thr Leu Ser Pro
165 170 175

Thr Lys Arg Leu Val His Ala Ser Glu Gln Tyr Ser Lys His Ser Gln
180 185 190

Glu Leu Tyr Ala Thr Ala Thr Leu Gln Arg Pro Gly Ser Leu Ala Ala
195 200 205

Gly Ser Arg Ala Ser Tyr Ser Ser Gln His Gly His Leu Gly Pro Glu
210 215 220

Leu Arg Ala Leu Gln Ser Pro Glu His His Ile Asp Pro Ile Tyr Glu
225 230 235 240

Asp Arg Val Tyr Gln Lys Pro Pro Met Arg Ser Leu Ser Gln Ser Gln
245 250 255

Gly Asp Pro Leu Pro Pro Ala His Thr Gly Thr Tyr Arg Thr Ser Thr
260 265 270

Ala Pro Ser Ser Pro Gly Val Asp Ser Val Pro Leu Gln Arg Thr Gly
275 280 285

Ser Gln His Gly Pro Gln Asn Ala Ala Ala Ala Thr Phe Gln Arg Ala
290 295 300

Ser Tyr Ala Ala Gly Pro Ala Ser Asn Tyr Ala Asp Pro Tyr Arg Gln
305 310 315 320

Leu Gln Tyr Cys Pro Ser Val Glu Ser Pro Tyr Ser Lys Ser Gly Pro
325 330 335

Ala Leu Pro Pro Glu Gly Thr Leu Ala Arg Ser Pro Ser Ile Asp Ser
340 345 350

Ile Gln Lys Asp Pro Arg Glu Phe Gly Trp Arg Asp Pro Glu Leu Pro
355 360 365

Glu Val Ile Gln Met Leu Gln His Gln Phe Pro Ser Val Gln Ser Asn
370 375 380

Ala Ala Ala Tyr Leu Gln His Leu Cys Phe Gly Asp Asn Lys Ile Lys
385 390 395 400

Ala Glu Ile Arg Arg Gln Gly Gly Ile Gln Leu Leu Val Asp Leu Leu
405 410 415

Asp His Arg Met Thr Glu Val His Arg Ser Ala Cys Gly Ala Leu Arg
420 425 430

Asn Leu Val Tyr Gly Lys Ala Asn Asp Asp Asn Lys Ile Ala Leu Lys
435 440 445

Asn Cys Gly Gly Ile Pro Ala Leu Val Arg Leu Leu Arg Lys Thr Thr
450 455 460

Asp Leu Glu Ile Arg Glu Leu Val Thr Gly Val Leu Trp Asn Leu Ser
465 470 475 480

Ser Cys Asp Ala Leu Lys Met Pro Ile Ile Gln Asp Ala Leu Ala Val
485 490 495

Leu Thr Asn Ala Val Ile Ile Pro His Ser Gly Trp Glu Asn Ser Pro
500 505 510

Leu Gln Asp Asp Arg Lys Ile Gln Leu His Ser Ser Gln Val Leu Arg
515 520 525

Asn Ala Thr Gly Cys Leu Arg Asn Val Ser Ser Ala Gly Glu Glu Ala
530 535 540

Arg Arg Arg Met Arg Glu Cys Asp Gly Leu Thr Asp Ala Leu Leu Tyr
545 550 555 560

Val Ile Gln Ser Ala Leu Gly Ser Ser Glu Ile Asp Ser Lys Thr Val
565 570 575

Glu Asn Cys Val Cys Ile Leu Arg Asn Leu Ser Tyr Arg Leu Ala Ala
580 585 590

Glu Thr Ser Gln Gly Gln His Met Gly Thr Asp Glu Leu Asp Gly Leu
595 600 605

Leu Cys Gly Glu Ala Asn Gly Lys Asp Ala Glu Ser Ser Gly Cys Trp
610 615 620

Gly Lys Lys Lys Lys Lys Lys Lys Ser Gln Asp Gln Trp Asp Gly Val
625 630 635 640

Gly Pro Leu Pro Asp Cys Ala Glu Pro Pro Lys Gly Ile Gln Met Leu
645 650 655

Trp His Pro Ser Ile Val Lys Pro Tyr Leu Thr Leu Leu Ser Glu Cys
660 665 670

Ser Asn Pro Asp Thr Leu Glu Gly Ala Ala Gly Ala Leu Gln Asn Leu
675 680 685

Ala Ala Gly Ser Trp Lys Trp Ser Val Tyr Ile Arg Ala Ala Val Arg
690 695 700

Lys Glu Lys Gly Leu Pro Ile Leu Val Glu Leu Leu Arg Ile Asp Asn
705 710 715 720

Asp Arg Val Val Cys Ala Val Ala Thr Ala Leu Arg Asn Met Ala Leu
725 730 735

Asp Val Arg Asn Lys Glu Leu Ile Gly Lys Tyr Ala Met Arg Asp Leu
740 745 750

Val His Arg Leu Pro Gly Gly Asn Asn Ser Asn Asn Thr Ala Ser Lys
755 760 765

Ala Met Ser Asp Asp Thr Val Thr Ala Val Cys Cys Thr Leu His Glu
770 775 780

Val Ile Thr Lys Asn Met Glu Asn Ala Lys Ala Leu Arg Asp Ala Gly
785 790 795 800

Gly Ile Glu Lys Leu Val Gly Ile Ser Lys Ser Lys Gly Asp Lys His
805 810 815

Ser Pro Lys Val Val Lys Ala Ala Ser Gln Val Leu Asn Ser Met Trp
820 825 830

Gln Tyr Arg Asp Leu Arg Ser Leu Tyr Lys Lys Asp Gly Trp Ser Gln
835 840 845

Tyr His Phe Val Ala Ser Ser Ser Thr Ile Glu Arg Asp Arg Gln Arg
850 855 860

Pro Tyr Ser Ser Ser Arg Thr Pro Ser Ile Ser Pro Val Arg Val Ser
865 870 875 880

Pro Asn Asn Arg Ser Ala Ser Ala Pro Ala Ser Pro Arg Glu Met Ile
885 890 895

Ser Leu Lys Glu Arg Lys Thr Asp Tyr Glu Cys Thr Gly Ser Asn Ala
900 905 910

Thr Tyr His Gly Gly Lys Gly Glu His Thr Ser Arg Lys Asp Ala Met
915 920 925

```
Thr Ala Gln Asn Thr Gly Ile Ser Thr Leu Tyr Arg Asn Ser Tyr Gly
    930                 935                 940

Ala Pro Ala Glu Asp Ile Lys His Asn Gln Val Ser Ala Gln Pro Val
945                 950                 955                 960

Pro Gln Glu Pro Ser Arg Lys Asp Tyr Glu Thr Tyr Gln Pro Phe Gln
                965                 970                 975

Asn Ser Thr Arg Asn Tyr Asp Glu Ser Phe Phe Glu Asp Gln Val His
            980                 985                 990

His Arg Pro Pro Ala Ser Glu Tyr Thr Met His Leu Gly Leu Lys Ser
        995                1000                1005

Thr Gly Asn Tyr Val Asp Phe Tyr Ser Ala Ala Arg Pro Tyr Ser Glu
    1010                1015                1020

Leu Asn Tyr Glu Thr Ser His Tyr Pro Ala Ser Pro Asp Ser Trp Val
1025                1030                1035                1040
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3907 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 142..3777
(D) OTHER INFORMATION: /note= "p0071"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
CTACTGTTGT TTTTGAGGGG CGGGCAGCCG CGCCGCCGCG GCACTTTTTT AATTTTTTCG      60

GGTGCCGCAG CAGCGACCCC TCGGCGCCGA TGTCCCTGAT CCCTGGAGCG ACGACGGCCG     120

CTGCCTAAGC TGGGAAGAGG A ATG CCA GCT CCT GAG CAG GCC TCA TTG GTG        171
                       Met Pro Ala Pro Glu Gln Ala Ser Leu Val
                         1               5                  10

GAG GAG GGG CAA CCA CAG ACC CGC CAG GAA GCT GCC TCC ACT GGC CCA        219
Glu Glu Gly Gln Pro Gln Thr Arg Gln Glu Ala Ala Ser Thr Gly Pro
                    15                  20                  25

GGC ATG GAA CCC GAG ACC ACA GCC ACC ACT ATT CTA GCA TCC GTG AAG        267
Gly Met Glu Pro Glu Thr Thr Ala Thr Thr Ile Leu Ala Ser Val Lys
                30                  35                  40

GAG CAG GAG CTT CAG TTT CAG CGA CTC ACC CGA GAA CTG GAA GTG GAA        315
Glu Gln Glu Leu Gln Phe Gln Arg Leu Thr Arg Glu Leu Glu Val Glu
            45                  50                  55

AGG CAG ATT GTT GCC AGT CAG CTA GAA AGA TGT AGG CTT GGA GCA GAA        363
Arg Gln Ile Val Ala Ser Gln Leu Glu Arg Cys Arg Leu Gly Ala Glu
        60                  65                  70

TCA CCA AGC ATC GCC AGC ACC AGC TCA ACT GAG AAG TCA TTT CCT TGG        411
Ser Pro Ser Ile Ala Ser Thr Ser Ser Thr Glu Lys Ser Phe Pro Trp
 75                  80                  85                  90

AGA TCA ACA GAC GTG CCA AAT ACT GGT GTA AGC AAA CCT AGA GTT TCT        459
Arg Ser Thr Asp Val Pro Asn Thr Gly Val Ser Lys Pro Arg Val Ser
                95                 100                 105

GAC GCT GTC CAG CCC AAC AAC TAT CTC ATC AGG ACA GAG CCA GAA CAA        507
Asp Ala Val Gln Pro Asn Asn Tyr Leu Ile Arg Thr Glu Pro Glu Gln
            110                 115                 120

GGA ACC CTC TAT TCA CCA GAA CAG ACA TCT CTC CAT GAA AGT GAG GGA        555
Gly Thr Leu Tyr Ser Pro Glu Gln Thr Ser Leu His Glu Ser Glu Gly
        125                 130                 135

TCA TTG GGT AAC TCA AGA AGT TCA ACA CAA ATG AAT TCT TAT TCC GAC        603
Ser Leu Gly Asn Ser Arg Ser Ser Thr Gln Met Asn Ser Tyr Ser Asp
    140                 145                 150

AGT GGA TAC CAG GAA GCA GGG AGT TTC CAC AAC AGC CAG AAC GTG AGC        651
Ser Gly Tyr Gln Glu Ala Gly Ser Phe His Asn Ser Gln Asn Val Ser
155                 160                 165                 170

AAG GCA GAC AAC AGA CAG CAG CAT TCA TTC ATA GGA TCA ACT AAC AAC        699
Lys Ala Asp Asn Arg Gln Gln His Ser Phe Ile Gly Ser Thr Asn Asn
                175                 180                 185

CAT GTG GTG AGG AAT TCA AGA GCT GAA GGA CAA ACA CTG GTT CAG CCA        747
His Val Val Arg Asn Ser Arg Ala Glu Gly Gln Thr Leu Val Gln Pro
```

```
                    190                 195                 200
TCA GTA GCC AAT CGG GCC ATG AGA AGA GTT AGT TCA GTT CCA TCT AGA      795
Ser Val Ala Asn Arg Ala Met Arg Arg Val Ser Ser Val Pro Ser Arg
        205                 210                 215

GCA CAG TCT CCT TCT TAT GTT ATC AGC ACA GGC GTG TCT CCT TCA AGG      843
Ala Gln Ser Pro Ser Tyr Val Ile Ser Thr Gly Val Ser Pro Ser Arg
    220                 225                 230

GGG TCT CTG AGA ACT TCT CTG GGT AGT GGA TTT GGC TCT CCG TCA GTG      891
Gly Ser Leu Arg Thr Ser Leu Gly Ser Gly Phe Gly Ser Pro Ser Val
235                 240                 245                 250

ACC GAC CCC CGA CCT CTG AAC CCC AGT GCA TAT TCC TCC ACC ACA TTA      939
Thr Asp Pro Arg Pro Leu Asn Pro Ser Ala Tyr Ser Ser Thr Thr Leu
                255                 260                 265

CCT GCT GCA CGG GCA GCC TCT CCG TAC TCA CAG AGA CCC GCC TCC CCA      987
Pro Ala Ala Arg Ala Ala Ser Pro Tyr Ser Gln Arg Pro Ala Ser Pro
            270                 275                 280

ACA GCT ATA CGG CGG ATT GGG TCA GTC ACC TCC CGG CAG ACC TCC AAT     1035
Thr Ala Ile Arg Arg Ile Gly Ser Val Thr Ser Arg Gln Thr Ser Asn
        285                 290                 295

CCC AAC GGA CCA ACC CCT CAA TAC CAA ACC ACC GCC AGA GTG GGG TCC     1083
Pro Asn Gly Pro Thr Pro Gln Tyr Gln Thr Thr Ala Arg Val Gly Ser
    300                 305                 310

CCA CTG ACC CTG ACG GAT GCA CAG ACT CGA GTA GCT TCC CCA TCC CAA     1131
Pro Leu Thr Leu Thr Asp Ala Gln Thr Arg Val Ala Ser Pro Ser Gln
315                 320                 325                 330

GGC CAG GTG GGG TCG TCG TCC CCC AAA CGC TCA GGG ATG ACC GCC GTA     1179
Gly Gln Val Gly Ser Ser Ser Pro Lys Arg Ser Gly Met Thr Ala Val
                335                 340                 345

CCA CAG CAT CTG GGA CCT TCA CTG CAA AGG ACT GTT CAT GAC ATG GAG     1227
Pro Gln His Leu Gly Pro Ser Leu Gln Arg Thr Val His Asp Met Glu
            350                 355                 360

CAA TTC GGA CAG CAG CAG TAT GAC ATT TAT GAG AGG ATG GTT CCA CCC     1275
Gln Phe Gly Gln Gln Gln Tyr Asp Ile Tyr Glu Arg Met Val Pro Pro
        365                 370                 375

AGG CCA GAC AGC CTG ACA GGC TTA CGG AGT TCC TAT GCT AGT CAG CAT     1323
Arg Pro Asp Ser Leu Thr Gly Leu Arg Ser Ser Tyr Ala Ser Gln His
    380                 385                 390

AGT CAG CTT GGG CAA GAC CTT CGT TCT GCC GTG TCT CCC GAC TTG CAC     1371
Ser Gln Leu Gly Gln Asp Leu Arg Ser Ala Val Ser Pro Asp Leu His
395                 400                 405                 410

ATT ACT CCT ATA TAT GAG GGG AGG ACC TAT TAC AGC CCA GTG TAC CGC     1419
Ile Thr Pro Ile Tyr Glu Gly Arg Thr Tyr Tyr Ser Pro Val Tyr Arg
                415                 420                 425

AGC CCA AAC CAT GGA ACT GTG GAG CTC CAA GGA TCG CAG ACG GCG TTG     1467
Ser Pro Asn His Gly Thr Val Glu Leu Gln Gly Ser Gln Thr Ala Leu
            430                 435                 440

TAT CGC ACA GGT GTA TCA GGT ATT GGA AAT CTA CAA AGG ACA TCC AGC     1515
Tyr Arg Thr Gly Val Ser Gly Ile Gly Asn Leu Gln Arg Thr Ser Ser
        445                 450                 455

CAA CGA AGT ACC CTT ACA TAC CAA AGA AAT AAT TAT GCT CTG AAC ACA     1563
Gln Arg Ser Thr Leu Thr Tyr Gln Arg Asn Asn Tyr Ala Leu Asn Thr
    460                 465                 470

ACA GCT ACC TAC GCG GAG CCC TAC AGG CCT ATA CAA TAC CGA GTG CAA     1611
Thr Ala Thr Tyr Ala Glu Pro Tyr Arg Pro Ile Gln Tyr Arg Val Gln
475                 480                 485                 490

GAG TGC AAT TAT AAC AGG CTT CAG CAT GCA GTG CCG GCT GAT GAT GGC     1659
Glu Cys Asn Tyr Asn Arg Leu Gln His Ala Val Pro Ala Asp Asp Gly
                495                 500                 505

ACC ACA AGA TCC CCA TCA ATA GAC AGC ATT CAG AAG GAC CCC AGG GAG     1707
Thr Thr Arg Ser Pro Ser Ile Asp Ser Ile Gln Lys Asp Pro Arg Glu
            510                 515                 520

TTT GCC TGG CGT GAT CCT GAG TTG CCT GAG GTC ATT CAC ATG CTT GAG     1755
Phe Ala Trp Arg Asp Pro Glu Leu Pro Glu Val Ile His Met Leu Glu
        525                 530                 535

CAC CAG TTC CCA TCT GTT CAG GCA AAT GCA GCG GCC TAC CTG CAG CAC     1803
His Gln Phe Pro Ser Val Gln Ala Asn Ala Ala Ala Tyr Leu Gln His
    540                 545                 550

CTG TGC TTT GGT GAC AAC AAA GTG AAG ATG GAG GTG TGT AGG TTA GGG     1851
Leu Cys Phe Gly Asp Asn Lys Val Lys Met Glu Val Cys Arg Leu Gly
555                 560                 565                 570

GGA ATC AAG CAT CTG GTT GAC CTT GT. GAT CAC AGA GTT TTG GAA GTT     1899
```

```
Gly Ile Lys His Leu Val Asp Leu Leu Asp His Arg Val Leu Glu Val
                575                 580                 585

CAG AAG AAT GCT TGT GGT GCC CTT CGA AAC CTC GTT TTT GGC AAG TCT      1947
Gln Lys Asn Ala Cys Gly Ala Leu Arg Asn Leu Val Phe Gly Lys Ser
            590                 595                 600

ACA GAT GAA AAT AAA ATA GCA ATG AAG AAT GTT GGT GGG ATA CCT GCC      1995
Thr Asp Glu Asn Lys Ile Ala Met Lys Asn Val Gly Gly Ile Pro Ala
        605                 610                 615

TTG TTG CGA CTG TTG AGA AAA TCT ATT GAT GCA GAA GTA AGG GAG CTT      2043
Leu Leu Arg Leu Leu Arg Lys Ser Ile Asp Ala Glu Val Arg Glu Leu
    620                 625                 630

GTT ACA GGA GTT CTT TGG AAT TTA TCC TCA TGT GAT GCT GTA AAA ATG      2091
Val Thr Gly Val Leu Trp Asn Leu Ser Ser Cys Asp Ala Val Lys Met
635                 640                 645                 650

ACA ATC ATT CGA GAT GCT CTC TCA ACC TTA ACA AAC ACT GTG ATT GTT      2139
Thr Ile Ile Arg Asp Ala Leu Ser Thr Leu Thr Asn Thr Val Ile Val
                655                 660                 665

CCA CAT TCT GGA TGG AAT AAC TCT TCT TTT GAT GAT GAT CAT AAA ATT      2187
Pro His Ser Gly Trp Asn Asn Ser Ser Phe Asp Asp Asp His Lys Ile
            670                 675                 680

AAA TTT CAG ACT TCA CTA GTT CTG CGT AAC ACG ACA GGT TGC CTA AGG      2235
Lys Phe Gln Thr Ser Leu Val Leu Arg Asn Thr Thr Gly Cys Leu Arg
        685                 690                 695

AAC CTC ACG TCC GCG GGG GAA GAA GCT CGG AAG CAA ATG CGG TCC TGC      2283
Asn Leu Thr Ser Ala Gly Glu Glu Ala Arg Lys Gln Met Arg Ser Cys
    700                 705                 710

GAG GGG CTG GTA GAC TCA CTG TTG TAT GTG ATC CAC ACG TGT GTG AAC      2331
Glu Gly Leu Val Asp Ser Leu Leu Tyr Val Ile His Thr Cys Val Asn
715                 720                 725                 730

ACA TCC GAT TAC GAC AGC AAG ACG GTG GAG AAC TGC GTG TGC ACC CTG      2379
Thr Ser Asp Tyr Asp Ser Lys Thr Val Glu Asn Cys Val Cys Thr Leu
                735                 740                 745

AGG AAC CTG TCC TAT CGG CTG GAG CTG GAG GTG CCC CAG GCC CGG TTA      2427
Arg Asn Leu Ser Tyr Arg Leu Glu Leu Glu Val Pro Gln Ala Arg Leu
            750                 755                 760

CTG GGA CTG AAC GAA TTG GAT GAC TTA CTA GGA AAA GAG TCT CCC AGC      2475
Leu Gly Leu Asn Glu Leu Asp Asp Leu Leu Gly Lys Glu Ser Pro Ser
        765                 770                 775

AAA GAC TCT GAG CCA AGT TGC TGG GGG AAG AAG AAG AAA AAG AAA AAG      2523
Lys Asp Ser Glu Pro Ser Cys Trp Gly Lys Lys Lys Lys Lys Lys Lys
    780                 785                 790

AGG ACT CCG CAA GAA GAT CAA TGG GAT GGA GTT GGT CCT ATC CCA GGA      2571
Arg Thr Pro Gln Glu Asp Gln Trp Asp Gly Val Gly Pro Ile Pro Gly
795                 800                 805                 810

CTG TCG AAG TCC CCC AAA GGG GTT GAG ATG CTG TGG CAC CCA TCG GTG      2619
Leu Ser Lys Ser Pro Lys Gly Val Glu Met Leu Trp His Pro Ser Val
                815                 820                 825

GTA AAA CCA TAT CTG ACT CTT CTA GCA GAA AGT TCC AAC CCA GCC ACC      2667
Val Lys Pro Tyr Leu Thr Leu Leu Ala Glu Ser Ser Asn Pro Ala Thr
            830                 835                 840

TTG GAA GGC TCT GCA GGG TCT CTC CAG AAC CTC TCT GCT AGC AAC TGG      2715
Leu Glu Gly Ser Ala Gly Ser Leu Gln Asn Leu Ser Ala Ser Asn Trp
        845                 850                 855

AAG TTT GCA GCA TAT ATC CGG GGC GGC CGT CCG AAA AGA AAA GGG CTC      2763
Lys Phe Ala Ala Tyr Ile Arg Gly Gly Arg Pro Lys Arg Lys Gly Leu
    860                 865                 870

CCC ATC CTT GTG GAG CTT CTG AGA ATG GAT AAC GAT AGA GTT GTT TCT      2811
Pro Ile Leu Val Glu Leu Leu Arg Met Asp Asn Asp Arg Val Val Ser
875                 880                 885                 890

TCC GGT GCA ACA GCC TTG AGG AAT ATG GCA CTA GAT GTT CGC AAC AAG      2859
Ser Gly Ala Thr Ala Leu Arg Asn Met Ala Leu Asp Val Arg Asn Lys
                895                 900                 905

GAG CTC ATA GGC AAA TAC GCC ATG CGA GAC CTG GTC AAC CGG CTC CCC      2907
Glu Leu Ile Gly Lys Tyr Ala Met Arg Asp Leu Val Asn Arg Leu Pro
            910                 915                 920

GGC GGC AAT GGC CCC AGT GTC TTG TCT GAT GAG ACC ATG GCA GCC ATC      2955
Gly Gly Asn Gly Pro Ser Val Leu Ser Asp Glu Thr Met Ala Ala Ile
        925                 930                 935

TGC TGT GCT CTG CAC GAG GTC ACC AGT AAA AAC ATG GAG AAC GCA AAA      3003
Cys Cys Ala Leu His Glu Val Thr Ser Lys Asn Met Glu Asn Ala Lys
    940                 945                 950
```

```
GCC CTG GCC GAC TCA GGA GGC ATA GAG AAG CTG GTG AAC ATA ACC AAA      3051
Ala Leu Ala Asp Ser Gly Gly Ile Glu Lys Leu Val Asn Ile Thr Lys
955                 960                 965                 970

GGC AGG GGC GAC AGA TCA TCT CTG AAA GTG GTG AAG GCA GCA GCC CAG      3099
Gly Arg Gly Asp Arg Ser Ser Leu Lys Val Val Lys Ala Ala Ala Gln
                975                 980                 985

GTC TTG AAT ACA TTA TGG CAA TAT CGG GAC CTG CGG AGC ATT TAT AAA      3147
Val Leu Asn Thr Leu Trp Gln Tyr Arg Asp Leu Arg Ser Ile Tyr Lys
            990                 995                1000

AAG GAT GGG TGG AAT CAG AAC CAT TTT ATT ACA CCT GTG TCG ACA TTG      3195
Lys Asp Gly Trp Asn Gln Asn His Phe Ile Thr Pro Val Ser Thr Leu
        1005                1010                1015

GAG CGA GAC CGA TTC AAA TCA CAT CCT TCC TTG TCT ACC ACC AAC CAA      3243
Glu Arg Asp Arg Phe Lys Ser His Pro Ser Leu Ser Thr Thr Asn Gln
    1020                1025                1030

CAG ATG TCA CCC ATC ATT CAG TCA GTC GGC AGC ACC TCT TCC TCA CCA      3291
Gln Met Ser Pro Ile Ile Gln Ser Val Gly Ser Thr Ser Ser Ser Pro
1035                1040                1045                1050

GCA CTG TTA GGA ATC AGA GAC CCT CGC TCT GAA TAC GAT AGG ACC CAG      3339
Ala Leu Leu Gly Ile Arg Asp Pro Arg Ser Glu Tyr Asp Arg Thr Gln
                1055                1060                1065

CCA CCT ATG CAG TAT TAC AAT AGC CAA GGG GAT GCC ACA CAT AAA GGC      3387
Pro Pro Met Gln Tyr Tyr Asn Ser Gln Gly Asp Ala Thr His Lys Gly
            1070                1075                1080

CTG TAC CCT GGC TCC AGC AAA CCT TCA CCA ATT TAC ATC AGT TCC TAT      3435
Leu Tyr Pro Gly Ser Ser Lys Pro Ser Pro Ile Tyr Ile Ser Ser Tyr
        1085                1090                1095

TCC TCA CCA GCA AGA GAA CAA AAT AGA CGG CTA CAG CAT CAA CAG CTG      3483
Ser Ser Pro Ala Arg Glu Gln Asn Arg Arg Leu Gln His Gln Gln Leu
    1100                1105                1110

TAT TAT AGT CAA GAT GAC TCC AAC AGA AAG AAC TTT GAT GCA TAC AGA      3531
Tyr Tyr Ser Gln Asp Asp Ser Asn Arg Lys Asn Phe Asp Ala Tyr Arg
1115                1120                1125                1130

TTG TAT TTG CAG TCT CCT CAT AGC TAT GAA GAT CCT TAT TTT GAT GAC      3579
Leu Tyr Leu Gln Ser Pro His Ser Tyr Glu Asp Pro Tyr Phe Asp Asp
                1135                1140                1145

CGA GTT CAC TTT CCA GCT TCT ACT GAT TAC TCA ACA CAG TAT GGA CTG      3627
Arg Val His Phe Pro Ala Ser Thr Asp Tyr Ser Thr Gln Tyr Gly Leu
            1150                1155                1160

AAA TCG ACC ACA AAT TAT GTA GAC TTT TAT TCC ACT AAA CGA CCT TCT      3675
Lys Ser Thr Thr Asn Tyr Val Asp Phe Tyr Ser Thr Lys Arg Pro Ser
        1165                1170                1175

TAT AGA GCA GAA CAG TAC CCA GGG TCC CCA GAC TCA TGG GTG TAC GAT      3723
Tyr Arg Ala Glu Gln Tyr Pro Gly Ser Pro Asp Ser Trp Val Tyr Asp
    1180                1185                1190

CAA GAT GCC CAA CAG AGG AAC TCT TTC TTT CTA ACC TTG TTC AGA TTG      3771
Gln Asp Ala Gln Gln Arg Asn Ser Phe Phe Leu Thr Leu Phe Arg Leu
1195                1200                1205                1210

AGG TGA AAAGTCCATC TTGCTGATTT CATGATTGAA ATGTGAAAGT GAAGTGGAAG      3827
Arg *

GAATGAATGA AGTGTGTTTT TTTTTCCTTT TTGAGGAATT ATCAGGGGAA TTCGATATCA   3887

AGCTTATCGA TACCGTCGAC                                               3907
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1212 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Pro Ala Pro Glu Gln Ala Ser Leu Val Glu Glu Gly Gln Pro Gln
 1               5                  10                  15

Thr Arg Gln Glu Ala Ala Ser Thr Gly Pro Gly Met Glu Pro Glu Thr
            20                  25                  30

Thr Ala Thr Thr Ile Leu Ala Ser Val Lys Glu Gln Glu Leu Gln Phe
        35                  40                  45

Gln Arg Leu Thr Arg Glu Leu Glu Val Glu Arg Gln Ile Val Ala Ser
    50                  55                  60
```

Gln Leu Glu Arg Cys Arg Leu Gly Ala Glu Ser Pro Ser Ile Ala Ser
65 70 75 80

Thr Ser Ser Thr Glu Lys Ser Phe Pro Trp Arg Ser Thr Asp Val Pro
85 90 95

Asn Thr Gly Val Ser Lys Pro Arg Val Ser Asp Ala Val Gln Pro Asn
100 105 110

Asn Tyr Leu Ile Arg Thr Glu Pro Glu Gln Gly Thr Leu Tyr Ser Pro
115 120 125

Glu Gln Thr Ser Leu His Glu Ser Glu Gly Ser Leu Gly Asn Ser Arg
130 135 140

Ser Ser Thr Gln Met Asn Ser Tyr Ser Asp Ser Gly Tyr Gln Glu Ala
145 150 155 160

Gly Ser Phe His Asn Ser Gln Asn Val Ser Lys Ala Asp Asn Arg Gln
165 170 175

Gln His Ser Phe Ile Gly Ser Thr Asn Asn His Val Val Arg Asn Ser
180 185 190

Arg Ala Glu Gly Gln Thr Leu Val Gln Pro Ser Val Ala Asn Arg Ala
195 200 205

Met Arg Arg Val Ser Ser Val Pro Ser Arg Ala Gln Ser Pro Ser Tyr
210 215 220

Val Ile Ser Thr Gly Val Ser Pro Ser Arg Gly Ser Leu Arg Thr Ser
225 230 235 240

Leu Gly Ser Gly Phe Gly Ser Pro Ser Val Thr Asp Pro Arg Pro Leu
245 250 255

Asn Pro Ser Ala Tyr Ser Ser Thr Thr Leu Pro Ala Ala Arg Ala Ala
260 265 270

Ser Pro Tyr Ser Gln Arg Pro Ala Ser Pro Thr Ala Ile Arg Arg Ile
275 280 285

Gly Ser Val Thr Ser Arg Gln Thr Ser Asn Pro Asn Gly Pro Thr Pro
290 295 300

Gln Tyr Gln Thr Thr Ala Arg Val Gly Ser Pro Leu Thr Leu Thr Asp
305 310 315 320

Ala Gln Thr Arg Val Ala Ser Pro Ser Gln Gly Gln Val Gly Ser Ser
325 330 335

Ser Pro Lys Arg Ser Gly Met Thr Ala Val Pro Gln His Leu Gly Pro
340 345 350

Ser Leu Gln Arg Thr Val His Asp Met Glu Gln Phe Gly Gln Gln Gln
355 360 365

Tyr Asp Ile Tyr Glu Arg Met Val Pro Pro Arg Pro Asp Ser Leu Thr
370 375 380

Gly Leu Arg Ser Ser Tyr Ala Ser Gln His Ser Gln Leu Gly Gln Asp
385 390 395 400

Leu Arg Ser Ala Val Ser Pro Asp Leu His Ile Thr Pro Ile Tyr Glu
405 410 415

Gly Arg Thr Tyr Tyr Ser Pro Val Tyr Arg Ser Pro Asn His Gly Thr
420 425 430

Val Glu Leu Gln Gly Ser Gln Thr Ala Leu Tyr Arg Thr Gly Val Ser
435 440 445

Gly Ile Gly Asn Leu Gln Arg Thr Ser Ser Gln Arg Ser Thr Leu Thr
450 455 460

Tyr Gln Arg Asn Asn Tyr Ala Leu Asn Thr Thr Ala Thr Tyr Ala Glu
465 470 475 480

Pro Tyr Arg Pro Ile Gln Tyr Arg Val Gln Glu Cys Asn Tyr Asn Arg
485 490 495

Leu Gln His Ala Val Pro Ala Asp Asp Gly Thr Thr Arg Ser Pro Ser
500 505 510

Ile Asp Ser Ile Gln Lys Asp Pro Arg Glu Phe Ala Trp Arg Asp Pro
515 520 525

Glu Leu Pro Glu Val Ile His Met Leu Glu His Gln Phe Pro Ser Val
530 535 540

Gln Ala Asn Ala Ala Ala Tyr Leu Gln His Leu Cys Phe Gly Asp Asn
545 550 555 560

Lys Val Lys Met Glu Val Cys Arg Leu Gly Gly Ile Lys His Leu Val

565 570 575

Asp Leu Leu Asp His Arg Val Leu Glu Val Gln Lys Asn Ala Cys Gly
580 585 590

Ala Leu Arg Asn Leu Val Phe Gly Lys Ser Thr Asp Glu Asn Lys Ile
595 600 605

Ala Met Lys Asn Val Gly Gly Ile Pro Ala Leu Leu Arg Leu Leu Arg
610 615 620

Lys Ser Ile Asp Ala Glu Val Arg Glu Leu Val Thr Gly Val Leu Trp
625 630 635 640

Asn Leu Ser Ser Cys Asp Ala Val Lys Met Thr Ile Ile Arg Asp Ala
645 650 655

Leu Ser Thr Leu Thr Asn Thr Val Ile Val Pro His Ser Gly Trp Asn
660 665 670

Asn Ser Ser Phe Asp Asp Asp His Lys Ile Lys Phe Gln Thr Ser Leu
675 680 685

Val Leu Arg Asn Thr Thr Gly Cys Leu Arg Asn Leu Thr Ser Ala Gly
690 695 700

Glu Glu Ala Arg Lys Gln Met Arg Ser Cys Glu Gly Leu Val Asp Ser
705 710 715 720

Leu Leu Tyr Val Ile His Thr Cys Val Asn Thr Ser Asp Tyr Asp Ser
725 730 735

Lys Thr Val Glu Asn Cys Val Cys Thr Leu Arg Asn Leu Ser Tyr Arg
740 745 750

Leu Glu Leu Glu Val Pro Gln Ala Arg Leu Leu Gly Leu Asn Glu Leu
755 760 765

Asp Asp Leu Leu Gly Lys Glu Ser Pro Ser Lys Asp Ser Glu Pro Ser
770 775 780

Cys Trp Gly Lys Lys Lys Lys Lys Lys Lys Arg Thr Pro Gln Glu Asp
785 790 795 800

Gln Trp Asp Gly Val Gly Pro Ile Pro Gly Leu Ser Lys Ser Pro Lys
805 810 815

Gly Val Glu Met Leu Trp His Pro Ser Val Val Lys Pro Tyr Leu Thr
820 825 830

Leu Leu Ala Glu Ser Ser Asn Pro Ala Thr Leu Glu Gly Ser Ala Gly
835 840 845

Ser Leu Gln Asn Leu Ser Ala Ser Asn Trp Lys Phe Ala Ala Tyr Ile
850 855 860

Arg Gly Gly Arg Pro Lys Arg Lys Gly Leu Pro Ile Leu Val Glu Leu
865 870 875 880

Leu Arg Met Asp Asn Asp Arg Val Val Ser Ser Gly Ala Thr Ala Leu
885 890 895

Arg Asn Met Ala Leu Asp Val Arg Asn Lys Glu Leu Ile Gly Lys Tyr
900 905 910

Ala Met Arg Asp Leu Val Asn Arg Leu Pro Gly Gly Asn Gly Pro Ser
915 920 925

Val Leu Ser Asp Glu Thr Met Ala Ala Ile Cys Cys Ala Leu His Glu
930 935 940

Val Thr Ser Lys Asn Met Glu Asn Ala Lys Ala Leu Ala Asp Ser Gly
945 950 955 960

Gly Ile Glu Lys Leu Val Asn Ile Thr Lys Gly Arg Gly Asp Arg Ser
965 970 975

Ser Leu Lys Val Val Lys Ala Ala Ala Gln Val Leu Asn Thr Leu Trp
980 985 990

Gln Tyr Arg Asp Leu Arg Ser Ile Tyr Lys Lys Asp Gly Trp Asn Gln
995 1000 1005

Asn His Phe Ile Thr Pro Val Ser Thr Leu Glu Arg Asp Arg Phe Lys
1010 1015 1020

Ser His Pro Ser Leu Ser Thr Thr Asn Gln Gln Met Ser Pro Ile Ile
1025 1030 1035 1040

Gln Ser Val Gly Ser Thr Ser Ser Ser Pro Ala Leu Leu Gly Ile Arg
1045 1050 1055

Asp Pro Arg Ser Glu Tyr Asp Arg Thr Gln Pro Pro Met Gln Tyr Tyr
1060 1065 1070

```
Asn Ser Gln Gly Asp Ala Thr His Lys Gly Leu Tyr Pro Gly Ser Ser
        1075                1080                1085

Lys Pro Ser Pro Ile Tyr Ile Ser Ser Tyr Ser Ser Pro Ala Arg Glu
    1090                1095                1100

Gln Asn Arg Arg Leu Gln His Gln Gln Leu Tyr Tyr Ser Gln Asp Asp
1105                1110                1115                1120

Ser Asn Arg Lys Asn Phe Asp Ala Tyr Arg Leu Tyr Leu Gln Ser Pro
                1125                1130                1135

His Ser Tyr Glu Asp Pro Tyr Phe Asp Asp Arg Val His Phe Pro Ala
            1140                1145                1150

Ser Thr Asp Tyr Ser Thr Gln Tyr Gly Leu Lys Ser Thr Thr Asn Tyr
        1155                1160                1165

Val Asp Phe Tyr Ser Thr Lys Arg Pro Ser Tyr Arg Ala Glu Gln Tyr
    1170                1175                1180

Pro Gly Ser Pro Asp Ser Trp Val Tyr Asp Gln Asp Ala Gln Gln Arg
1185                1190                1195                1200

Asn Ser Phe Phe Leu Thr Leu Phe Arg Leu Arg *
                1205                1210
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 970 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..970
(D) OTHER INFORMATION: /note= "Y2H9"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
GAATTCCCAC AGATACCACT GCTGCTCCCG CCCTTTCGCT CCTCGGCCGC GCAATGGGCA      60
CCCGCGACGA CGAGTACGAC TACCTCTTTA AAGTTGTCCT TATTGGAGAT TCTGGTGTTG     120
GAAAGAGTAA TCTCCTGTCT CGATTTACTC GAAATGAGTT TAATCTGGAA AGCAAGAGCA     180
CCATTGGAGT AGAGTTTGCA ACAAGAAGCA TCCAGGTTGA TGGAAAAACA ATAAAGGCAC     240
AGATATGGGA CACAGCAGGG CAAGAGCGAT ATCGAGCTAT AACATCAGCA TATTATCGTG     300
GAGCTGTAGG TGCCTTATTG GTTTATGACA TTGCTAAACA TCTCACATAT GAAAATGTAG     360
AGCGATGGCT GAAAGAACTG AGAGATCATG CTGATAGTAA CATTGTTATC ATGCTTGTGG     420
GCAATAAGAG TGATCTACGT CATCTCAGGG CAGTTCCTAC AGATGAAGCA AGAGCTTTTG     480
CAGAAAAGAA TGGTTTGTCA TTCATTGAAA CTTCGGCCCT AGACTCTACA AATGTAGAAG     540
CTGCTTTTCA GACAATTTTA ACAGAGATTT ACCGCATTGT TTCTCAGAAG CAAATGTCAG     600
ACAGACGCGA AAATGACATG TCTCCAAGCA ACAATGTGGT TCCTATTCAT GTTCCACCAA     660
CCACTGAAAA CAAGCCAAAG GTGCAGTGCT GTCAGAACAT CTAAGGCATT TCTCTTCTCC     720
CCTAGAAGGC TGTGTATAGT CCATTTCCCA GGTCTSASAT TTAAATATAW TTGTAATTCT     780
TGTGTCACTT TTGTGTTTTA TTACTTCATA CTTATGAATT TTTCCATGTC CTAAGTCTTT     840
TGATTTTGMT TTATAAAATC ATCCACTTGT NCCGAATGNC TGCAGCTTTT TTTCATGCTA     900
TGGCTTCACT AGCCTTAGTT TNATAAACTG AATGTTTGGA TTCCTCCCCC CAAAAAAAAA     960
AAAACTCGAG                                                           970
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 264 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..264
(D) OTHER INFORMATION: /note= "Y2H23b"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GAATTCGCGG CCGNGTCGAC CCCCCACCCC CGATGCCACC ACCCCCANTG GGNTCTCCCN 60

NCCCAGTCAT CAGTTCTTCC ATGGNGTNCC CTGGTCTGCC CCCTCCAGCT CCCCCAGGCN 120

TTCTCCGGGT CTGNCAGCAG CCNCCAGATT AACTCAACAG TGTCACTCCC TGGGGGTGGG 180

TCTGGNCCCC CTGANGATGT GAAGCCACCA GTCTNAGNGG TCCGGGGTCT GTACTGTCCA 240

CCCCCTCCAG GTGGACCTGG CGCT 264

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 340 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..340
(D) OTHER INFORMATION: /note= "Y2H27"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GAATTCGCGG CCGCGTCGAC CGCGGTCGCG TCGACCTGTT GCCCAGGCCC TAGAGGTCAT 60

TCCTCGTACC CTGATCCAGA ACTGTGGGGC CAGCACCATC CGTCTACTTA CCTCCCTTCG 120

GGCCAAGCAC ACCCAGGAGA ACTGTGAGAC CTGGGGTGTA AATGGTGAGA CGGGTACTTT 180

GGTGGACATG AAGGAACTGG GCATATGGGA GCCATTGGCT GTGAAGCTGC AGACTTATAA 240

GACAGCAGTG GAGACGGCAG TTCTGCTACT GCGAATTGAT GACATCGTTT CAGGCCACAA 300

AAAGAAAGGC GATGACCAGA GCCGGCAAGG CGGNGCTCCT 340

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 404 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..404
(D) OTHER INFORMATION: /note= "Y2H35"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAATTCGCGG TCGCGTCGAC GGTTAGTCCC ACTGGNCGCA TCGAGGGNTT CACCAACGTC

ATGGAGCTGT ATGGCANGAT CGCCGAGGTC TTCCNCCTGC CAACTGCCGA GGTGATGTTC

TGCACCCTGA NCACCCACAA AGTGGACATN GACAAGCTCC TGGGGGGCCA GATCGGGCTG

GAGGACTTCA TCTTCGCCCA CGTGAAGGGG YAGCGCAAGG AGGTGGAGGT GTTCAWGTCG

GAGGATGYAC TCGGKCTCAC CATCACGGAC AACGGGGCTG GCTACGCTTC CATCAAGCGC

ATCAAGGAGG GCAGCGTGAT CGACCACATC CACCTCATCA GCGTGGGCGA CATGATCGAG

GCCATTAACG GGCAGAGCTT CCTGGGCTGC CGGCATTACG AGGT

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 350 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..350
(D) OTHER INFORMATION: /note= "Y2H171"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GAATTCGCGG CCGCGTCGAC AAAAAAAGTA AAAGGAACTC GGCAAATCTT ACCCCGCCTG

TTTACCAAAA ACATCACCTC TAGCATCACC AGTATTAGAG GCACCGCCTG CCCAGTGACA

CATGTTTAAC GGCCGCGGTA CCCTAACCGT GCAAAGGTAG CATAATCACT TGTTCCTTAA 100

GTAGGGACCT GTATGAATGG CTCCACGAGG GTTCAGCTGT CTCTTACTTT TAACCAATGA 150

AATTGACCTG CCCGTGAAGA GGCGGGCATG ACACAGCAAG ACGAGAAGAC CCTATGGAGC 200

TTTAATTTAT TAATGCAAAC AGTACCTAAC AAACCCACAG GGTCCTAAAC 350

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 350 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 1..350
(D) OTHER INFORMATION: /note= "Y2H41"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GAATTCGCGG NCGCGTCGAC AGATAATGAA AAAACCAGAG GTTCCCTTCT TTGGTCCCCT 60

NNNNGATGGT GCTATTGTGA ATGGAAAGGT TCTACCCATT ATGGTTAGAG CAACAGCTAT 120

AAATGCAAGC CGTGCTCTGA AATCTCTGAT TCCATTGTAT CAAAACTTCT ATGAGGAGAG 180

AGCACGATAC CTGCAAACAA TTGTCCAGCA CCACTTAGAA CCAACAACAT TTGAAGATTT 240

TGNAGCACAG GTTTTTTCTC CAGCTCCCTA CCACCATTTA CCATCTGATG CCGTTGGCTC 300

CTACCCAGAG ATTCTACCCA GTGAAAACTC CCACAGCAAC GCAGGTAGGA 350

CLAIMS

What is claimed is:

1. An isolated nucleic acid comprising a nucleotide sequence encoding at least a presenilin-interacting domain of a presenilin-interacting protein selected from the group consisting of a mammalian S5a (approximately residues 70-377 of SEQ ID NO: 2), GT24 (approximately residues 346-862 of SEQ ID NO: 4), p0071 (approximately residues 509-1022 of SEQ ID NO: 6), Rab11 (SEQ ID NO: 7), retinoid X receptor-β (SEQ ID NO:8), cytoplasmic chaperonin (SEQ ID NO: 9), Y2H35 (SEQ ID NO: 10), Y2H171 (SEQ ID NO: 11), and a Y2H41 (SEQ ID NO: 12) presenilin-interacting domain.

2. An isolated nucleic acid comprising a nucleotide sequence of at least 10 consecutive nucleotides selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, GenBank Accession Numbers F08730, T18858, X81889, X56740, X53143, M84820, X63522, M81766, U17104, X74801, R12984, D55326, and T64843, and a sequence complementary to any of these sequences.

3. An isolated nucleic acid as in claim 2 comprising a nucleotide sequence of at least 15 consecutive nucleotides selected from said group.

4. An isolated nucleic acid as in claim 2 comprising a nucleotide sequence of at least 20 consecutive nucleotides selected from said group.

5. An isolated nucleic acid comprising a nucleotide sequence encoding an antigenic determinant of a presenilin-interacting protein selected from the group consisting of a mammalian S5a, GT24, p0071, Rab11, retinoid X receptor-β, cytoplasmic chaperonin, Y2H35, Y2H171, and Y2H41 protein.

6. A method for identifying allelic variants or heterospecific homologues of a human presenilin-interacting protein gene comprising

choosing a nucleic acid probe or primer capable of hybridizing to a human presenilin-interacting protein gene sequence under stringent hybridization conditions;

mixing said probe or primer with a sample of nucleic acids which may contain a nucleic acid corresponding to said variant or homologue;

detecting hybridization of said probe or primer to said nucleic acid corresponding to said variant or homologue.

7. A method as in claim 6 wherein said sample comprises a sample of nucleic acids selected from the group consisting of human genomic DNA, human mRNA, and human cDNA.

8. A method as in claim 6 wherein said sample comprises a sample of nucleic acids selected from the group consisting of mammalian genomic DNA, mammalian mRNA, and mammalian cDNA.

9. A method as in claim 6 wherein said sample comprises a sample of nucleic acids selected from the group consisting of invertebrate genomic DNA, invertebrate mRNA, and invertebrate cDNA.

10. A method as in claim 6 further comprising the step of isolating said nucleic acid corresponding to said variant or homologue.

11. A method as in claim 6 wherein said nucleic acid is identified by hybridization.

12. A method as in claim 6 wherein said nucleic acid is identified by PCR amplification.

13. A method for identifying allelic variants or heterospecific homologues of a human presenilin-interacting protein gene comprising

choosing an antibody capable of selectively binding to a human presenilin-interacting protein;

mixing said antibody with a sample of proteins which may contain a protein corresponding to said variant or homologue;

detecting binding of said antibody to said protein corresponding to said variant or homologue.

14. A method as in claim 13 wherein said sample comprises a sample of proteins selected from the group consisting of human proteins, human fusion proteins, and proteolytic fragments thereof.

15. A method as in claim 13 wherein said sample comprises a sample of proteins selected from the group consisting of mammalian proteins, mammalian fusion proteins, and proteolytic fragments thereof.

16. A method as in claim 13 wherein said sample comprises a sample of proteins selected from the group consisting of invertebrate proteins, invertebrate fusion proteins, and proteolytic fragments thereof.

17. A method as in claim 13 further comprising the step of substantially purifying said protein corresponding to said variant or homologue.

18. An isolated nucleic acid comprising an allelic variant or a heterospecific homologue of a human presenilin-interacting protein gene.

19. An isolated nucleic acid encoding an allelic variant or heterospecific homologue of a human presenilin-interacting protein.

20. An isolated nucleic acid comprising a recombinant vector including a nucleotide sequence of any one of claims 1-19.

21. An isolated nucleic acid as in claim 20 wherein said vector is an expression vector and said presenilin-interacting protein nucleotide sequence is operably joined to a regulatory region.

22. An isolated nucleic acid as in claim 21 wherein said expression vector may express said presenilin-interacting protein sequence in mammalian cells.

23. An isolated nucleic acid as in claim 22 wherein said cells are selected from the group consisting of fibroblast, liver, kidney, spleen, bone marrow and neurological cells.

24. An isolated nucleic acid as in claim 21 wherein said vector is selected from the group consisting of vaccinia virus, adenovirus, retrovirus, neurotropic viruses and Herpes simplex.

25. An isolated nucleic acid as in claim 21 wherein said expression vector encodes at least a presenilin-interacting domain of a presenilin-interacting protein selected from the group consisting of a mammalian S5a, GT24, p0071, Rab11, retinoid X receptor-β, cytoplasmic chaperonin, Y2H35, Y2H171, and Y2H41 protein.

26. An isolated nucleic acid as in claim 21 wherein said vector further comprises sequences encoding an exogenous protein operably joined to said presenilin-interacting protein sequence and whereby said vector encodes a presenilin-interacting protein fusion protein.

27. An isolated nucleic acid as in claim 26 wherein said exogenous protein is selected from the group consisting of lacZ, trpE, maltose-binding protein, a poly-His

tag, glutathione-S-transferase, a GAL4-DNA binding domain, and a GAL4 activation domain.

28. An isolated nucleic acid comprising a recombinant expression vector including nucleotide sequences corresponding to an endogenous regulatory region of a presenilin-interacting protein gene.

29. An isolated nucleic acid as in claim 28 wherein said endogenous regulatory region is operably joined to a marker gene.

30. A host cell transformed with an expression vector of any one of claims 20-29, or a descendant thereof.

31. A host cell as in claim 30 wherein said host cell is selected from the group consisting of bacterial cells and yeast cells.

32. A host cell as in claim 30 wherein said host cell is selected from the group consisting of fetal cells, embryonic stem cells, zygotes, gametes, and germ line cells.

33. A host cell as in claim 30 wherein said cell is selected from the group consisting of fibroblast, liver, kidney, spleen, bone marrow and neurological cells.

34. A host cell as in claim 30 wherein said cell is an invertebrate cell.

35. A non-human animal model for Alzheimer's Disease, wherein a genome of said animal, or an ancestor thereof, has been modified by at least one recombinant construct, and wherein said recombinant construct has introduced a modification selected from the group consisting of (1) insertion of nucleotide sequences encoding at least a functional domain of a heterospecific normal presenilin-interacting protein, (2) insertion of nucleotide sequences encoding at least a functional domain of a heterospecific mutant presenilin-interacting protein, (3) insertion of nucleotide sequences encoding at least a functional domain of a conspecific homologue of a heterospecific mutant presenilin-interacting protein, and (4) inactivation of an endogenous presenilin-interacting protein gene.

36. An animal as in claim 35 wherein said modification is insertion of a nucleotide sequence encoding at least a functional domain of a normal human presenilin-interacting protein selected from the group consisting of a mammalian S5a, GT24, p0071, Rab11, retinoid X receptor-β, cytoplasmic chaperonin, Y2H35, Y2H171, and Y2H41 protein.

37. An animal as in claim 35 wherein said modification is insertion of a nucleotide sequence encoding at least a functional domain of a mutant human presenilin-interacting protein selected from the group consisting of a mammalian S5a, GT24, p0071, Rab11, retinoid X receptor-β, cytoplasmic chaperonin, Y2H35, Y2H171, and Y2H41 protein.

38. An animal as in claim 35 wherein said animal is selected from the group consisting of rats, mice, hamsters, guinea pigs, rabbits, dogs, cats, goats, sheep, pigs, and non-human primates.

39. An animal as in claim 35 wherein said animal is an invertebrate.

40. A method for producing at least a functional domain of a presenilin-interacting protein comprising culturing a host cell of any of claims 30-34 under suitable conditions to produce said presenilin by expressing said nucleic acid.

41. A substantially pure preparation of a protein selected from the group consisting of a mammalian S5a, GT24, p0071, Rab11, retinoid X receptor-β, cytoplasmic chaperonin, Y2H35, Y2H171, and Y2H41 protein.

42. A substantially pure preparation of a polypeptide comprising an amino acid sequence of at least 10 consecutive amino acid residues selected from the group consisting SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 6, and GenBank Accession Numbers F08730, T18858, X81889, X56740, X53143, M84820, X63522, M81766, U17104, X74801, R12984, D55326, and T64843.

43. A substantially pure preparation of a polypeptide as in claim 42 comprising an amino acid sequence of at least 15 consecutive amino acid residues selected from said group.

44. A substantially pure preparation of a polypeptide comprising at least a presenilin-interacting domain of a presenilin-interacting protein selected from the group consisting of a mammalian S5a, GT24, p0071, Rab11, retinoid X receptor-β, cytoplasmic chaperonin, Y2H35, Y2H171, and Y2H41 protein.

45. A substantially pure preparation of a polypeptide comprising an antigenic determinant of a presenilin-interacting protein selected from the group consisting of a mammalian S5a, GT24, p0071, Rab11, retinoid X receptor-β, cytoplasmic chaperonin, Y2H35, Y2H171, and Y2H41 protein.

46. A method of producing antibodies which selectively bind to a presenilin-interacting protein comprising the steps of

administering an immunogenically effective amount of a presenilin-interacting protein immunogen to an animal;

allowing said animal to produce antibodies to said immunogen; and

obtaining said antibodies from said animal or from a cell culture derived therefrom.

47. A substantially pure preparation of an antibody which selectively binds to an antigenic determinant of a presenilin-interacting protein selected from the group consisting of a mammalian S5a, GT24, p0071, Rab11, retinoid X receptor-β, cytoplasmic chaperonin, Y2H35, Y2H171, and Y2H41 protein.

48. A substantially pure preparation of an antibody as in claim 47 wherein said antibody selectively binds to an antigenic determinant of a mutant presenilin-interacting protein and fails to bind to a normal presenilin-interacting protein.

49. A cell line producing an antibody of any one of claims 47-48.

50. A method for identifying compounds which can modulate the expression of a presenilin-interacting protein gene comprising

contacting a cell with a test candidate wherein said cell includes a regulatory region of a presenilin-interacting protein gene operably joined to a coding region; and

detecting a change in expression of said coding region.

51. A method as in claim 50 wherein said change comprises a change in a level of an mRNA transcript encoded by said coding region.

52. A method as in claim 50 wherein said change comprises a change in a level of a protein encoded by said coding region.

53. A method as in claim 50 wherein said change is a result of an activity of a protein encoded by said coding region.

54. A method as in claim 50 wherein said coding region encodes a marker protein selected from the group consisting of β-galactosidase, alkaline phosphatase, green fluorescent protein, and luciferase.

55. A method for identifying compounds which can selectively bind to a presenilin-interacting protein comprising the steps of

providing a preparation including at least one presenilin-interacting protein component;

contacting said preparation with a sample including at least one candidate compound; and

detecting binding of said presenilin-interacting protein component to said candidate compound.

56. The method in 55 wherein said binding to said presenilin-interacting component is detected by an assay selected from the group consisting of: affinity chromatography, co-immunoprecipitation, a Biomolecular Interaction Assay, and a yeast two-hybrid system.

57. A method of identifying compounds which can modulate activity of a presenilin-interacting protein comprising the steps of

providing a cell expressing a normal or mutant presenilin-interacting protein gene;

contacting said cell with at least one candidate compound; and

detecting a change in a marker of said activity.

58. A method as in claim 57 wherein measurement of said marker indicates a difference between cells bearing an expressed mutant presenilin-interacting protein gene and otherwise identical cells free of an expressed mutant presenilin-interacting protein gene.

59. A method as in claim 57 wherein said change comprises a change in a non-specific marker of cell physiology selected from the group consisting of pH; intracellular Ca^{++}, Na^{+}, or K^{+}; cyclic AMP levels; GTP/GDP ratios; phosphatidylinositol activity; and protein phosphorylation.

60. A method as in claim 57 wherein said change comprises a change in expression of said presenilin-interacting protein.

61. A method as in claim 57 wherein said change comprises a change in intracellular concentration or flux of an ion selected from the group consisting of Ca^{2+}, Na^{+} and K^{+}.

62. A method as in claim 57 wherein said change comprises a change in occurrence or rate of apoptosis or cell death.

63. A method as in claim 57 wherein said change comprises a change in production of Aβ peptides.

64. A method as in claim 57 wherein said change comprises a change in phosphorylation of at least one microtubule associated protein.

65. A method as in claim 57 wherein said cell is a cell cultured in vitro.

66. A method as in claim 65 wherein said cell is a transformed host cell of any one of claims 30-34.

67. A method as in claim 65 wherein said cell is explanted from a host bearing at least one mutant presenilin-interacting protein gene.

68. A method as in claim 65 wherein said cell is explanted from a transgenic animal of any one of claims 35-39.

69. A method as in claim 57 wherein said cell is a cell in a live animal.

70. A method as in claim 69 wherein said cell is a cell of a transgenic animal of any one of claims 35-39.

71. A method as in claim 57 wherein said cell is in a human subject in a clinical trial.

72. A diagnostic method for determining if a subject bears a mutant presenilin-interacting protein gene comprising the steps of

providing a biological sample of said subject;

detecting in said sample a mutant presenilin-interacting protein nucleic acid, a mutant presenilin-interacting protein, or a mutant presenilin-interacting protein activity.

73. A method as in claim 72, wherein a mutant presenilin-interacting protein nucleic acid is detected by an assay selected from the group consisting of direct nucleotide sequencing, probe specific hybridization, restriction enzyme digest and mapping, PCR mapping, ligase-mediated PCR detection, RNase protection, electrophoretic mobility shift detection, and chemical mismatch cleavage.

74. A method as in claim 72, wherein a mutant presenilin-interacting protein is detected by an assay selected from the group consisting of an immunoassay, a protease assay, and an electrophoretic mobility assay.

75. A pharmaceutical preparation comprising a substantially pure presenilin-interacting protein and a pharmaceutically acceptable carrier.

76. A pharmaceutical preparation comprising an expression vector operably encoding a presenilin-interacting protein, wherein said expression vector may express said presenilin-interacting protein in a human subject, and a pharmaceutically acceptable carrier.

77. A pharmaceutical preparation comprising an expression vector operably encoding a presenilin-interacting protein antisense sequence, wherein said expression vector may express said presenilin-interacting protein antisense sequence in a human subject, and a pharmaceutically acceptable carrier.

78. A pharmaceutical preparation comprising a substantially pure antibody, wherein said antibody selectively binds to a mutant presenilin-interacting protein, and a pharmaceutically acceptable carrier.

79. A pharmaceutical preparation as in claim 78 wherein said preparation is essentially free of an antibody which selectively binds a normal presenilin-interacting protein.

80. A pharmaceutical preparation comprising a substantially pure preparation of an antigenic determinant of a mutant presenilin-interacting protein.

81. A pharmaceutical preparation as in claim 80 wherein said preparation is essentially free of an antigenic determinant of a normal presenilin-interacting protein.

82. A method of treatment for a patient bearing a mutant presenilin-interacting protein gene comprising the step of administering to said patient a therapeutically effective amount of the pharmaceutical preparation of any one of claims 75-81.

83. A method as in claim 82, wherein said pharmaceutical preparation is targeted to a cell type is selected from the group consisting of heart, brain, lung, liver, skeletal muscle, kidney, pancreas and neurological cells.

5

INTERNATIONAL SEARCH REPORT

Int ional Application No
PCT/CA 97/00051

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C12N15/74 C12N5/10 C07K14/47 C07K16/18
A01K67/027 C12Q1/68 G01N33/577 A61K31/70 A61K38/17
A61K39/395 A61K48/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K A01K C12N G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 270, no. 35, 1 September 1995, pages 20615-20620, XP002029380 JOHANSSON, E., ET AL . ON: "MOLECULAR CLONING AND EXPRESSION OF A PITUITARY GLAND PROTEIN MODULATING INTESTINAL FLUID SECRETION" cited in the application see the whole document --- -/--	1-12, 18-20, 30,31, 34, 41-47,49

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
14 April 1997	25.04.97
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+ 31-70) 340-3016	Authorized officer Holtorf, S

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int ional Application No
PCT/CA 97/00051

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EMBL SEQUENCE DATA LIBRARY, 20 December 1994, HEIDELBERG, GERMANY, XP002029381 PAWLAK, A., ET AL . : "CHARACTERIZATION OF A LARGE POPULATION OF mRNAs FROM HUMAN TESTIS" cited in the application ACCESSION No. T18858 see the whole document	1-12, 18-20, 30,31,34
X	EMBL SEQUENCE DATA LIBRARY , 17 February 1995, HEIDELBERG, GERMANY, XP002029382 AUFFRAY, C. , ET AL . : "IMAGE: INTEGRATED MOLECULAR ANALYSIS OF THE HUMAN GENOME AND ITS EXPRESSION" cited in the application ACCESSION No. F08730 see the whole document	1-12, 18-20, 30,31, 34, 41-47,49
X	EMBL SEQUENCE DATA LIBRARY, 22 July 1994, HEIDELBERG, GERMANY, XP002029383 ZAHRAOUI, A., ET AL . : "CODING SEQUENCES OF HUMAN Rab8 AND Rab11 cDNAs" cited in the application ACCESSION No. X56740 see the whole document	1-12, 18-20, 30,31, 34, 41-47,49
X	EMBL SEQUENCE DATA LIBRARY , 19 February 1991, HEIDELBERG, GERMANY, XP002029384 DRIVAS, G.T., ET AL.: "IDENTIFICATION AND CHARACTERIZATION OF A HUMAN HOMOLOG OF THE SCHIZOSACCHAROMYCES POMBE RAS-LIKE GENE YPT-3" cited in the application ACCESSION No. X53143 see the whole document	1-12, 18-20, 30,31, 34, 41-47,49
X	EMBL SEQUENCE DATA LIBRARY, 7 February 1992, HEIDELBERG, GERMANY, XP002029385 CHAMBON,P., ET AL . : "PURIFICATION , CLONING , AND RXR IDENTITY OF THE HELA CELL FACTOR WITH WHICH RAR OR TR HETERODIMERIZES TO BIND TARGET SEQUENCES EFFICIENTLY" cited in the application ACCESSION No. M84820 see the whole document	1-12, 18-20, 30,31, 34, 41-47,49

-/--

INTERNATIONAL SEARCH REPORT

Inte ·onal Application No
PC·/CA 97/00051

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
X	EMBL SEQUENCE DATA LIBRARY, 31 March 1992, HEIDELBERG, GERMANY, XP002029386 FLEISCHHAUER, K., ET AL . : "ISOLATION OF A FULL-LENGTH cDNA CLONE ENCODING A N-TERMINALLY VARIANT FORM OF THE HUMAN RETINOID X RECEPTOR BETA" cited in the application ACCESSION No. X63522 see the whole document	1-12, 18-20, 30,31, 34, 41-47,49
X	EMBL SEQUENCE DATA LIBRARY, 10 December 1991, HEIDELBERG, GERMANY, XP002029387 YU, V.C., ET AL . : "RXR-beta: A COREGULATOR THAT ENHANCES BINDING OF RETINOIC ACID, THYROID HORMONE, AND VITAMIN D RECEPTORS TO THEIR COGNATE RESPONSE ELEMENTS" cited in the application ACCESSION No. M81766 see the whole document	1-12, 18-20, 30,31, 34, 41-47,49
X	EMBL SEQUENCE DATA LIBRARY, 7 January 1995, HEIDELBERG, GERMANY, XP002029388 SEVIGNY,G., ET AL.: "ASSIGNMENT OF THE HUMAN HOMOLOGUE OF THE mTRiC-P5 GENE (TRIC5) TO BAND 1q23 BY FLUORESCENCE IN SITU HYBRIDIZATION" cited in the application ACCESSION No. U17104 see the whole document	1-12, 18-20, 30,31, 34, 41-47,49
X	EMBL SEQUENCE DATA LIBRARY, 1 January 1994, HEIDELBERG, GERMANY, XP002029389 WALKLEY, N.A., ET AL . : "CLONING , STRUCTURE AND mRNA EXPRESSION OF HUMAN Cctg, WHICH ENCODES THE CHAPERONIN SUBUNIT CCTgamma" cited in the application ACCESSION No. X74801 see the whole document	1-12, 18-20, 30,31, 34, 41-47,49
X	EMBL SEQUENCE DATA LIBRARY, 22 April 1995, HEIDELBERG, GERMANY, XP002029390 HILLIER, L., ET AL . : "THE WashU-MERCK EST PROJECT" cited in the application ACCESSION No. R12984 see the whole document	1-12, 18-20, 30,31, 34, 41-47,49

-/--

Form PCT ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 97/00051

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EMBL SEQUENCE DATA LIBRARY, 25 August 1995, HEIDELBERG, GERMANY, XP002029391 FUJIWARA, T., ET AL .: ""UNPUBLISHED"" cited in the application ACCESSION No. D55326 see the whole document ---	1-12, 18-20, 30,31,34
X	EMBL SEQUENCE DATA LIBRARY, 6 March 1995, HEIDELBERG, GERMANY, XP002029392 HILLIER,L., ET AL . : "THE WashU-MERCK EST PROJECT" cited in the application ACCESSION No. T64843 see the whole document -----	1-12, 18-20, 30,31, 34, 41-47,49

1